Atto[redacted] Docket No.: 06501-056001

APPLICATION

FOR

UNITED STATES LETTERS PATENT

TITLE: PROTEIN HAVING PDZ DOMAIN SEQUENCE

APPLICANT: SHIN-ICHI FUNAHASHI AND SHOJI MIYATA

CERTIFICATE OF MAILING BY EXPRESS MAIL

Express Mail Label No. EL228025793US

I hereby certify under 37 CFR §1.10 that this correspondence is being deposited with the United States Postal Service as Express Mail Post Office to Addressee with sufficient postage on the date indicated below and is addressed to the Assistant Commissioner for Patents, Washington, D.C. 20231.

February 11, 2000
Date of Deposit

Paul Bea
Signature

Paul Bea
Typed or Printed Name of Person Signing Certificate

PROTEIN HAVING PDZ DOMAIN SEQUENCE

This is a continuation-in-part of PCT/JP98/03603 filed August 12, 1998, and claims priority from Japanese Patent Application Nos. 9/230356, filed August 12, 1997, and 10/189944, filed June 18, 1998.

Technical Field

The present invention relates to novel proteins having the PDZ domain sequence and also to gene encoding the proteins.

Background Art

Proteins such as PSD-95, hDlg, ZO-1, p55, Dsh, LIN-7, InaD, and PTPL1/FAP1 are known to possess the PDZ domain and are called the PDZ family. A structure having approximately 80 to 90 amino acid residues, repeated three times and each containing a conserved "Gly-Leu-Gly-Phe (GLGF)" 4 amino acid motif (Neuron 9:929-942 (1992)), was initially identified in the 95 KDa post-synaptic density protein, PSD-95. The same domain structure was later found in the Drosophila lethal (1) discs large-1 tumor suppressor protein, DlgA (Cell 66:451-464 (1991)), and in the tight junction protein, ZO-1 (J. Cell Biol. 121:491-502 (1993)). The repeat sequence was therefore named the "PDZ domain" by combining the initials of PSD-95, DlgA, and ZO-1. (It is also called the "GLGF repeat" or "DHR (DlgA homology region) domain.") A protein having the PDZ domain is known to bind to other proteins by means of the sequence of this PDZ domain. For example, the PSD-95 protein is known to bind to the NMDA receptor 2B (Kornau, H. C., et al., Science 269:1737-1740 (1995)) and the Shaker-type K^+ channel (Kim, E., et al., Nature 378:85-88 (1995)). The hDlg

protein has been reported to bind directly to the protein encoded by the adenomatous polyposis coli tumor suppressor gene/APC (Matsumine et al., Science 272:1020-1023 (1996)), and the Dsh protein has been reported to bind directly to the Notch protein (Axelrod, J. D., et al., Science 271:1826-1832 (1996)). Furthermore, the InaD protein has been reported to bind to a Ca^{2+} channel protein, TRP, that functions in the Drosophila visual signal transduction cascade (Shieh, B. H. and Zhu, M. Y., Neuron 16:991-998 (1996)). The structure of proteins having the PDZ domain varies because some of the proteins contain only one domain (p55 and Dsh), while others contain two (SIP-1: Yanagisawa, J., et al., J. Biol. Chem. 272:7167-7172 (1997)), three (PSD-95 and hDlg), five (InaD and PTPL1/FAP1), seven (GRIP: Dong, H., et al., Nature 386:279-284 (1997)), or thirteen (Ullmer, C., et al., FEBS Letters 424:63-68 (1998)). Also a recently reported mouse gene lacks a region encoding an N-terminal peptide of the protein, but which encodes a peptide having four PDZ domains within this incomplete genetic region (Recorded to GenBank on May 18, 1997; accession number AF000168). Although there are a few exceptions, proteins having the PDZ domain are known to bind to other proteins that have a hydrophobic amino acid region consisting of three amino acids represented by "Thr/Ser-Xaa-Val" (Xaa being an arbitrary amino acid residue) at their C-terminus. Most of these proteins are transmembrane proteins and are presumed to function in signal transduction within the cell (TIBS 21:455-458 (1996), Yanagisawa, J., et al., J. Biol. Chem. 272:7167-7172 (1997)).

Since the above proteins having the PDZ domain and proteins that interact with these proteins are involved in

neural transmission, apoptosis, and malignant conversion, they have recently drawn attention as targets for developing pharmaceuticals.

Disclosure of the Invention

An objective of the present invention is to provide a novel protein having the PDZ domain sequence and a DNA encoding the protein. Another objective of the present invention is to provide a vector containing the DNA, a transformant harboring the DNA in which the DNA can be expressed, and a method of producing the recombinant protein utilizing the transformant. A further objective of the present invention is to provide an antisense DNA against the DNA and antibody that binds to the protein. Still another objective of the present invention is to provide a screening method for proteins that bind to the PDZ-domain protein.

While analyzing the changes of gene expression in human umbilical vascular endothelial cells by TNFα, the present inventors isolated a gene whose expression was increased by TNFα stimulation. Screening was performed using the gene as a probe, and, as a result, a gene encoding novel proteins was isolated. The present inventors analyzed the structure of the proteins encoded by the isolated gene and found that the proteins contain within the molecule the PDZ domain sequence that plays an important role in the interactions with other proteins involved in neural transmission, apoptosis, and malignant conversion. The present inventors also found that the single gene produces at least five different transcriptional products through the differences in transcription initiation sites and in splicing.

The present inventors succeeded in preparing the proteins encoded by the gene as recombinant proteins by incorporating the isolated gene into an expression vector, and by transfecting it into *E. coli* cells and culturing the cells. In addition, by immunizing rabbits with the proteins thus prepared, the present inventors succeeded in preparing antibodies that bind to the proteins.

The present invention relates to a group of novel proteins having the PDZ domain sequence within the molecule and to their gene, and more specifically, to

(1) a protein comprising the amino acid sequence described in SEQ ID NOs: 1, 2, 82, 83, or 84;

(2) a protein comprising the amino acid sequence described in SEQ ID NOs: 1, 2, 82, 83, or 84, in which one or more amino acids have been substituted, deleted, and/or added, and having affinities to other proteins characteristic to the PDZ domain;

(3) a fusion protein comprising the protein described in (1) or (2) and a protein or a peptide containing at least one antibody recognition site;

(4) a DNA encoding the protein of any one of (1) through (3);

(5) an antisense DNA against the DNA or a part thereof whose nucleotide sequence is described in SEQ ID NO: 2;

(6) a vector containing the DNA of (4);

(7) a transformant harboring the DNA of (4), in which the DNA can be expressed;

(8) a method of producing the protein of any one of (1) through (3), comprising the process of culturing the transformants described in (7);

(9) a screening method for proteins that bind to the protein of (1) or (2), comprising the process of selecting the proteins that bind to the proteins by contacting sample proteins with the proteins of any one of (1) through (3);

(10) a screening method for genes encoding the proteins that bind to the proteins of (1) or (2), comprising the process of selecting the genes corresponding to the gene products that bind to the proteins of (1) or (2) by contacting the gene products of the sample genes with the protein of (1) or (2);

(11) a protein that binds to the protein of (1) or (2);

(12) the protein of (11) that can be isolated by the method of (9);

(13) a gene encoding a protein that bind to the protein of (1) or (2);

(14) the gene of (13) that can be isolated by the method of (10); and

(15) an antibody that bind to the protein of (1) or (2).

In the present invention, the "PDZ domain sequence" refers to a sequence having 80 to 90 amino acids, containing the four amino acid motif that consists of "Gly-Leu-Gly-Phe" or similar amino acids (cf. TIBS 20:102-103 (1995)).

The present invention relates to novel proteins having the PDZ domain sequence. Although there are a few exceptions, proteins having the PDZ domain are known to interact with other proteins that have a hydrophobic amino acid region at their C-terminal ends. The other proteins are transmembrane proteins and are presumed to function in signal transduction within the cell (TIBS 21: 455-458

(1996), Yanagisawa, J., et al., J. Biol. Chem. 272:7167-7172 (1997)).

The present inventors have discovered five different transcription products among those that encode proteins having the PDZ domain. These products are thought to arise from a single gene through differences in transcription initiation sites and in splicing. The amino acid sequences of the proteins encoded by these transcription products are shown in SEQ ID NOs: 1, 2, 82, 83, and 84.

The protein having the amino acid sequence described in SEQ ID NO: 1, which is included in the proteins of the present invention, possesses nine PDZ domains that correspond to amino acid positions 69 to 158 (SEQ ID NO: 4), positions 371 to 461 (SEQ ID NO: 5), positions 520 to 615 (SEQ ID NO: 6), positions 649 to 734 (SEQ ID NO: 7), positions 782 to 865 (SEQ ID NO: 8), positions 928 to 1013 (SEQ ID NO: 9), positions 1024 to 1108 (SEQ ID NO: 10), positions 1161 to 1249 (SEQ ID NO: 11), and positions 1286 to 1373 (SEQ ID NO: 12)(see Figure 8).

Similarly, the protein having the amino acid sequence described in SEQ ID NO: 2, which is also included in the proteins of the present invention, corresponds to amino acids 369 to 1373 of the sequence described in SEQ ID NO: 1. The difference between the structures of these proteins is considered to arise from the difference in the mRNA transcription initiation sites.

The protein described in SEQ ID NO: 2 possesses a total of eight PDZ domain sequences, corresponding to amino acids 3 to 93, 152 to 247, 281 to 366, 414 to 497, 560 to 645, 656 to 740, 793 to 881, and 918 to 1005. However, it does not possess the first PDZ domain found in the protein described in SEQ ID NO: 1. Although its biological

significance is not clear, considering the specific expression of the mRNA corresponding to the protein described in SEQ ID NO: 2 in the liver (Example 5) and the fact that the PDZ domain plays an important role in protein-protein interactions, the protein described in SEQ ID NO: 2, by lacking this domain, may be involved in controlling the signal in the liver differently from the other tissues.

The protein having the amino acid sequence described in SEQ ID NO: 82 (the 32-8-1a protein), which is also included in the proteins of the present invention, consists of 2,000 amino acids. These amino acids are predicted by combining sequences of two cDNAs. One cDNA was discovered in the search for a cDNA derived from the human brain and contains a 5' upstream region of a cDNA encoding a protein having the amino acid sequence described in SEQ ID NO: 1. The other cDNA (SEQ ID NO: 3) encodes the protein having the amino acid sequence described in SEQ ID NO: 1. The 32-8-1a protein possesses a total of 13 PDZ domain sequences, corresponding to amino acids 133 to 222, 253 to 335, 373 to 461, 549 to 632, 696 to 784, 1004 to 1087, 1147 to 1240, 1276 to 1361, 1409 to 1492, 1555 to 1640, 1651 to 1735, 1788 to 1870, and 1913 to 2000 (Figure 25).

Similarly, the amino acid sequences of the proteins encoded by the two splicing variants that are thought to result from the different splicing from the transcription product encoding the 32-8-1a protein and are also included in the proteins of the present invention, are shown in SEQ ID NO: 83 (the 32-8-1b protein) and in SEQ ID NO: 84 (the 32-8-1c protein). The 32-8-1b protein, similar to the 32-8-1a protein, consists of 2,070 amino acids, possessing 13 PDZ domains. The PDZ domains of the 32-8-1b protein exist

at positions 133 to 222, 253 to 335, 373 to 461, 549 to 632, 696 to 784, 1004 to 1087, 1147 to 1241, 1346 to 1431, 1479 to 1562, 1625 to 1710, 1721 to 1805, 1858 to 1946, and 1983 to 2070 of its amino acid sequence.

In contrast, the 32-8-1c protein has a shorter chain length than 32-8-1a or 32-8-1b because of the termination codon created by the splicing, and consists of 1,239 amino acids, possessing seven PDZ domains. The PDZ domains of the 32-8-1c protein exist at positions 133 to 222, 253 to 335, 373 to 461, 549 to 632, 696 to 784, 1004 to 1087, and 1147 to 1239 of its amino acid sequence.

It is clinically very significant that these proteins of the present invention are all of human origin, as opposed to being derived from other animals. In particular, proteins derived from other organisms (e.g., mice or rats) cause serious side effects such as reduction or loss of therapeutic effects by generating antibodies or by inducing serum sickness and anaphylactic shock, due to the immunogenicity when they are used to treat humans. Therefore, it is desirable to use proteins of human origin as therapeutic materials for humans.

The proteins of the present invention can be prepared from natural proteins, but they can also be prepared as recombinant proteins using recombinant genetics technology. The natural proteins can be isolated from such sources as the human umbilical vascular endothelial cells (HUVEC) by means of methods well-known to persons skilled in the art. For example, they can be isolated as described below, with an affinity column in which an antibody against the protein of the present invention has been bound to an appropriate support. The affinity column can be constructed, for example, according to the method described by Wilchek et

al. (Wilchek et al., Methods Enzymol. 104:3-55 (1984)). Furthermore, the recombinant protein can be prepared by culturing the cells transformed with the DNA encoding the protein of the present invention, as will be described later.

The proteins of the present invention also include functional derivatives of the proteins having the amino acid sequences described in SEQ ID NOs: 1, 2, 82, 83, and 84. A "functional derivative" means a protein that differs from the amino acid sequences described in SEQ ID NOs: 1, 2, 82, 83, and 84 by one or more amino acid residues through substitution, deletion, or addition, but that still maintains the affinity to the other proteins characteristic of the PDZ domain. This affinity normally arises from the affinity to a hydrophobic amino acid region that exists in the C-terminal ends of the other proteins. The hydrophobic amino acid region contains a hydrophobic amino acid motif represented by "Thr/Ser-Xaa-Val" (Xaa being an arbitrary amino acid residue) (cf. Science 269:1737 (1995), Nature 378:85 (1995), Science 277:1511 (1997), Neuron 20:693 (1998), Oncogene 16:643 (1998), J. Biol. Chem. 273:1591 (1998), Science 272:1020 (1996), Proc. Natl. Acad. Sci. USA 94:6670 (1997), Proc. Natl. Acad. Sci. USA 94:11612 (1997), J. Neurosci. 18:128 (1998), J. Neurosci. 16:7407 (1996), Nature Biotech. 15:336 (1997), FEBS Letters 409:53 (1997), Nature 386:284 (1997), Nature 386:279 (1997), Nature Structure Biol. 5:19 (1998), J. Neurosci. 16:24 (1996), J. Biol. Chem. 272:24191 (1997), Science 271:1826 (1996), TIBS 21:455 (1996), Cell 85:195 (1996), Neuron 18:95 (1997), Proc. Natl. Acad. Sci. USA 94:12682 (1997), J. Biol. Chem. 272:8539 (1997), J. Biol. Chem. 272:24333 (1997), J. Biol. Chem. 272:7167 (1997), Proc. Natl. Acad. Sci. USA 94:13683

(1997), Nature 392:6676 (1998), J. Biol. Chem. 272:32019 (1997), Mol. Biol. Cell 9:671 (1998)).

Functional derivatives occur naturally or can be produced artificially; both of these are included in the present invention. Methods to alter amino acids, which are well known to persons skilled in the art, include the methods developed by Kunkel et al. (Methods Enzymol. 85:2763-2766 (1988)) and those that utilize polymerase chain reaction (PCR). In the Kunkel method, uracil is incorporated by using dut$^-$ or ung$^-$ *E. coli* as a host when preparing the single-stranded DNA to be used as the template. Primers containing the desired mutations are annealed to this template containing uracil, and ordinary DNA synthesis is performed *in vitro*. When the double-stranded DNA thus produced with the uracil-containing DNA is introduced into ordinary *E. coli* cells, the uracil-containing DNA strand becomes degraded, and DNA synthesis proceeds with the mutated DNA strand as the template. As a result, DNA into which mutations have been introduced can be obtained with a very high efficiency. An example of the methods of introducing mutations using PCR follows. Two sets of primers are prepared. One of the primers in each set encompasses the region into which the mutation will be introduced, and the other contains a restriction enzyme recognition site or a sequence just outside of it. A region containing appropriate restriction enzyme sites is thus targeted. PCR reactions are then performed with the two sets of primers. After the products of the two PCR reactions are mixed, the DNA is amplified using primers having sequences corresponding to the recognition sites of the two restriction enzymes or the sequences just outside of them. The product is next digested with appropriate

restriction enzymes so that the resultant fragment contains the region into which the mutation has been introduced. The fragment thus obtained is substituted for the said region in the original DNA (Saiki et al., Science 239:487-491 (1988), Current Protocols in Molecular Biology, Greene Publishing Associates and Wiley - Interscience Publishing, Unit 8.5.1-8.5.10 (1997), Supplement to Jikken Igaku (Experimental Medicine) "Shin Idenshi Kogaku Handbook (New Genetic Engineering Handbook)," Yohdosha, pp251-261). The desired number of amino acids to be substituted in a functional derivative is generally 10 or less, more preferably 6 or less, and still more preferably 3 or less.

The term "substantially pure" as used herein in reference to a given polypeptide means that the polypeptide is substantially free from other biological macromolecules. The substantially pure polypeptide is at least 75% (e.g., at least 80, 85, 95, or 99%) pure by dry weight. Purity can be measured by any appropriate standard method, for example, by column chromatography, polyacrylamide gel electrophoresis, or HPLC analysis.

A "conservative amino acid substitution" is one in which an amino acid residue is replaced with another residue having a chemically similar side chain. Families of amino acid residues having similar side chains have been defined in the art. These families include amino acids with basic side chains (e.g., lysine, arginine, histidine), acidic side chains (e.g., aspartic acid, glutamic acid), uncharged polar side chains (e.g., glycine, asparagine, glutamine, serine, threonine, tyrosine, cysteine), nonpolar side chains (e.g., alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan), beta-branched side chains (e.g., threonine, valine, isoleucine)

and aromatic side chains (e.g., tyrosine, phenylalanine, tryptophan, histidine).

The present invention also relates to the DNA encoding the proteins of the present invention described above. The DNA encoding the proteins of the present invention can be cDNA, genomic DNA, or synthetic DNA. The DNA of the present invention can be used, for example, to produce the proteins of the present invention as recombinant proteins. More specifically, the proteins of the present invention can be prepared as recombinant proteins by inserting the DNA encoding the proteins of the present invention into appropriate expression vectors, culturing the transformants obtained by introducing the said vectors into appropriate cells, and purifying the expressed proteins.

By hybridization under "stringent conditions" is meant hybridization at 37°C, 1 X SSC, followed by washing at 42°C, 0.5 X SSC.

The "percent identity" of two amino acid sequences or of two nucleic acids is determined using the algorithm of Karlin and Altschul (Proc. Natl. Acad. Sci. USA 87:2264-2268, 1990), modified as in Karlin and Altschul (Proc. Natl. Acad. Sci. USA 90:5873-5877, 1993). Such an algorithm is incorporated into the NBLAST and XBLAST programs of Altschul et al. (J. Mol. Biol. 215:403-410, 1990). BLAST nucleotide searches are performed with the NBLAST program, score = 100, wordlength = 12. BLAST protein searches are performed with the XBLAST program, score = 50, wordlength = 3. Where gaps exist between two sequences, Gapped BLAST is utilized as described in Altschul et al. (Nucleic Acids Res. 25:3389-3402, 1997). When utilizing BLAST and Gapped BLAST programs, the default

parameters of the respective programs (e.g., XBLAST and NBLAST) are used. See http://www.ncbi.nlm.nih.gov.

An "isolated nucleic acid" is a nucleic acid the structure of which is not identical to that of any naturally occurring nucleic acid or to that of any fragment of a naturally occurring genomic nucleic acid spanning more than three separate genes. The term therefore covers, for example, (a) a DNA which has the sequence of part of a naturally occurring genomic DNA molecule but is not flanked by both of the coding sequences that flank that part of the molecule in the genome of the organism in which it naturally occurs; (b) a nucleic acid incorporated into a vector or into the genomic DNA of a prokaryote or eukaryote in a manner such that the resulting molecule is not identical to any naturally occurring vector or genomic DNA; (c) a separate molecule such as a cDNA, a genomic fragment, a fragment produced by polymerase chain reaction (PCR), or a restriction fragment; and (d) a recombinant nucleotide sequence that is part of a hybrid gene, i.e., a gene encoding a fusion protein. Specifically excluded from this definition are nucleic acids present in mixtures of different (i) DNA molecules, (ii) transfected cells, or (iii) cell clones: e.g., as these occur in a DNA library such as a cDNA or genomic DNA library.

The cells to be used for producing the recombinant proteins include, but are not limited to, animal cells such as Chinese hamster ovary (CHO) cells, COS cells (a cell line obtained by transforming monkey CV-1 fibroblasts by the SV40 virus lacking the replication origin), mouse NIH3T3 cells, human HeLa cells, and human lymphoid Namalva cells (Current Protocols in Molecular Biology, Greene Publishing Associates and Wiley - Interscience Publishing,

Unit 16.12-16.14 (1991)). As the vectors, pSV2neo, pcDNAI, pCD8, pRcRSV, pREP4, pCEP4 (Invitrogen), pMAM, pMAMneo (Clontech), pCI-neo mammalian expression vector, pSI-neo mammalian expression vector, pTARGET™ mammalian expression vector (Promega), and the like can be used. Both plasmid vectors and recombinant viruses can be constructed for producing the recombinant protein. Recombinant adenoviruses using the pAdex vector (Supplement to Jikken Igaku (Experimental Medicine) "Shin Idenshi Kogaku Handbook (New Genetic Engineering Handbook)," Yohdosha, pp238-244), the LN and LXSN vector series, the pBabe vector series (a modified version of the preceding series), recombinant retroviruses using such vectors as the MFG vectors (Supplement to Jikken Igaku (Experimental Medicine) "Shin Idenshi Kogaku Handbook (New Genetic Engineering Handbook)," Yohdosha, pp245-250, Current Protocols in Molecular Biology, Greene Publishing Associates and Wiley - Interscience Publishing, Unit 9.10.1-9.14.3 (1992)), Sindbis viruses, and vaccinia viruses (Current Protocols in Molecular Biology, Greene Publishing Associates and Wiley - Interscience Publishing, Unit 16.15.1-16.19.9 (1992)) can also be used to produce the recombinant proteins. It is also possible to produce the recombinant proteins by utilizing baculoviruses, and silkworm larvae. Alternatively, cultured cell lines such as SF21, SF9, and High Five™ cells can be used as the host (Supplement to Jikken Igaku (Experimental Medicine) "Bio Manual Series 7: Bunshi Seibutsu Kenkyu No Tame No Tanpakushitsu Jikken Hou (Protein Experimentation Methods for Molecular Biology Research)," Yohdosha, pp167-171 (1994), OReilly, D. R. et al., "Baculovirus Expression Vectors, A Laboratory Manual," Oxford University Press (1992)). As the baculovirus

expression vectors, pBacPAK8, 9, pBacPAK-His 1/2/3, pAcUW31 (Clontech), pBlueBac (Invitrogen), pBAC, pBACgus (Novagen), etc., can be used.

The promoters utilized to express the proteins efficiently in animal cells include, for example, the SV40 early promoter (Rigby In Williamson (ed.), Genetic Engineering, Vol. 3, Academic Press, London, pp83-141 (1982)), the EF-1α promoter (Kim et al., Gene 91:217-223 (1990)), the CAG promoter (Niwa et al., Gene 108:193-200 (1991)), the RSV LTR promoter (Cullen, Methods in Enzymology 152:684-704 (1987)), the SRα promoter (Takabe et al., Mol. Cell. Biol. 8:466 (1988)), the CMV early promoter (Seed and Aruffo, Proc. Natl. Acad. Sci. USA 84:3365-3369 (1987)), the SV40 late promoter (Gheysen and Fiers, J. Mol. Appl. Genet. 1:385-394 (1982)), the Adenovirus late promoter (Kaufman et al., Mol. Cell. Biol. 9:946 (1989)), the HSV TK promoter, and inducible expression promoters. The MMTV promoter induced by glucocorticoids, the MT (metallothionein) II promoter induced by phorbol esters or heavy metals, the Tet-On/Off system that can be turned on and off by tetracycline (Clontech), the expression system that can be induced by ecdysone (Invitrogen), and the Lac Switch expression system induced by IPTG are preferred examples of the inducible expression promoters.

It is also possible to use yeast cells to produce the proteins. Protease-deficient cell lines such as BJ2168, BJ926, and CB023, and cell lines for secretion vectors, such as 20B-12, can be used as hosts (Supplement to Jikken Igaku (Experimental Medicine) "Bio Manual Series 4: Idenshi Donyu To Hatsugen Kaisekihou (Gene Introduction and Expression Analysis Methods)," Yohdosha, pp166-176 (1994)). The expression vectors include pYEUra3 (Clontech), pYEX™-

BX, and pYEX™-S1. It is also possible to express the protein in fission yeast SP-Q01, using fission yeast expression vector pESP-1 (Stratagene). The PGK promoter and the ADH1 promoter, which are constitutive; the CUP1 promoter, which is inducible by copper ions; the Gal1-Gal10 promoter, which is induced by galactose and repressed by glucose; and the PHO5 promoter, which is induced by a reduction in phosphate concentrations and repressed by high phosphate concentrations are preferable as promoters that efficiently express the protein in the yeast cells. In fission yeast, promoters such as the nmt1 promoter are preferable.

Four broad categories of expression promoters can be used to produce recombinant proteins using *E. coli* cells. The λPL promoter is regulated by the clts857 repressor and is induced by heat shock. N4830-1 and M5219 can be used as the host, and vectors such as pPL-lambda, pKC30, and pRIT2T can be used for expression. The tac promoter is regulated by the $lacI^q$ repressor and is induced by adding isopropyl β-D-thiogalactoside (IPTG). JM105 and XL1-Blue can be used as the host, and vectors such as pDR540, pKK233-3, pGEX-3X, and pMAL-c2 can be used for expression. The trp promoter is regulated by the trp repressor and is induced by adding β indole acrylic acid (IAA). HB101 and the like can be used as the host; vectors such as pBTrp2 can be used for expression. The T7 phage promoter is recognized for expression by only the T7RNA polymerase. Therefore, the BL21(DE3) strain can be used as the host. This strain can be prepared by lysogenizing the *E. coli* BL21 strain with λ phage DE3, into which the lacI gene and a DNA fragment containing the T7RNA polymerase gene under the control of the lacUV5 promoter are inserted within its int gene. The

inducible expression directed by the T7 promoter becomes possible by adding IPTG, which induces the T7RNA polymerase. The vectors include pET-3c and pET-8c. BL21(DE3)pLysS contains, in addition to the above plasmid, a plasmid producing the T7 lysozyme. This lysozyme is a natural inhibitor that binds to the T7RNA polymerase and inhibits its transcription, in order to suppress the basal level T7RNA polymerase activity. Therefore, BL21(DE3)pLysS can also be used as the host. pET-11c, pET-11d, and the like, which possess the T7lac promoter with the lac operator sequence inserted downstream of the T7 promoter transcription initiation site, can also be used as the expression vector (Studier, F., et al., J. Mol. Biol. 189:113-130 (1996), Studier, F., et al., Methods Enzymol. 185:60-8 (1990)).

Methods of introducing the vector into the host include the electroporation method (Chu, G., et al., Nucl. Acids Res. 15:1311-1326 (1987)), the calcium phosphate method (Chen, C. and Okayama, H., Mol. Cell. Biol. 7:2745-2752 (1987)), the DEAE dextran method (Lopata, M. A., et al., Nucl. Acids Res. 12:5707-5717 (1984); Sussman, D. J. and Milman, G., Mol. Cell. Biol. 4:1642-1643 (1985)), and the lipofectin method (Derijard, B., Cell 7:1025-1037 (1994); Lamb, B. T., et al., Nature Genetics 5:22-30 (1993); Rabindran, S. K., et al., Science 259:230-234 (1993)), but any method can be used.

The recombinant protein can be purified from the transformant thus obtained by means of the gel filtration method, ion exchange chromatography, affinity chromatography, reverse phase chromatography, hydroxyapatite chromatography, hydrogen bonding chromatography, and chelating columns (Deutscher, M. P.,

ed., Methods Enzymol. 182, Guide to Protein Purification, 1990; Principles and Methods Series: Gel Filtration, Ion Exchange chromatography, and Affinity chromatography. Pharmacia). Antibodies against the protein of the present invention are prepared as described below, and the protein can be highly purified by means of affinity chromatography using the antibodies.

Persons skilled in the art can, by using the prepared protein of the present invention, easily prepare the antibodies that bind to it. The antibodies of the present invention can be obtained by expressing the gene of the present invention using an appropriate *E. coli* expression vector; purifying the product; and immunizing rabbits, mice, rats, goats, or chickens with it. It is also possible to synthesize peptides that correspond to appropriate regions of the protein encoded by the gene of the present invention, and to immunize the animals described above, thereby obtaining the antibodies to the gene product. Methods to establish mouse or rat hybridomas can be used to produce monoclonal antibodies (Kohler and Milstein, Nature 256:495-497 (1975)). Specifically, mice, rats, or Armenian hamsters are first immunized with the prepared protein of the present invention. The antibody-producing cells are then collected from the spleen or the lymph nodes and fused in vitro with myeloma cells, and clones are selected through screening using the antigen (Harlow, E. and Lane, D., Antibodies: A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York (1988)). The mouse myeloma cells include p3-x63-Ag8-U1 (P3-U1), P3-NSI/1-Ag4-1 (NS-1), and SP2/0-Ag14 (AP2/0), and the rat myeloma cells include YB2/3HL.P2G11.16Ag20 (YB2/0). The cells can be fused using polyethylene glycol or

electric pulses. Monoclonal antibodies, such as that contained in the cultured supernatant of the hybridomas and that contained in the ascites of the mouse treated with an immunosuppressant and with the mass-cultured hybridoma injected into its abdominal cavity, can be purified by, for example, protein A-Sepharose (Pharmacia). Furthermore, monoclonal antibodies can also be purified using an affinity column having the protein of the present invention immobilized onto the support (Harlow, E. and Lane, D., Antibodies: A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York (1988)).

When the antibodies thus obtained are administered to humans, it is beneficial to use a human or humanized antibody in order to reduce the immunogenicity. The methods to humanize antibodies include the CDR graft method, in which the antibody gene is cloned from the monoclonal antibody producing cells and the antigen determining region is transplanted to a known human antibody (Immunology Methods Manual 1: pp98-107, Academic Press). Human antibodies can also be produced by immunizing a mouse that has its immune system replaced with the human immune system, following a procedure similar to the one used with regular monoclonal antibodies. The human B cell hybridoma method (Kozbor, et al., Immunology Today 4:72 (1983)), and the Epstein-Barr virus (EBV) - Hybridoma method (Cole, et al. in Monoclonal Antibodies and Cancer Therapy, Ala R. Liss, Inc. pp77-96 (1985)) can also be used to produce monoclonal antibodies.

The antibodies thus obtained can be used not only to detect the proteins of the present invention and as antibody therapies, but also to screen the proteins

described below that interact with the proteins of the present invention.

The present invention also relates to the methods used to screen for the proteins that bind to the proteins of the present invention. The group of proteins having the PDZ domain, such as those of the present invention, share a common property of interacting with other proteins having the region of hydrophobic amino acids on the C-terminus. These and other binding proteins can be isolated by the screening methods of the present invention. These screening methods include the process to select the proteins that bind to the proteins of the present invention. In such a process, the sample proteins are brought into contact with the proteins of the present invention in the form of lysates from the cells or tissues that are expected to contain the target proteins.

An example of the specific methods is the immunoprecipitation method. The immunoprecipitation method is the most common method used to detect protein-protein binding. In immunoprecipitation, biological samples, such as lysates from cells or tissues, for example, cell lysates prepared by dissolving cells such as human umbilical vascular endothelial cells with Triton X-100 or sodium deoxycholate, are usually brought into contact with the proteins of the present invention. The antibodies are then applied to the complex thus formed between the proteins of the present invention with their binding proteins. The immune complexes thus formed are then precipitated (Supplement to Jikken Igaku (Experimental Medicine) "Shin Idenshi Kogaku Handbook (New Genetic Engineering Handbook)," Yohdosha, pp304-308 (1996)).

The immune complex can be precipitated by, for example, using protein A-Sepharose or protein G-Sepharose when the antibody is a mouse IgG antibody. General methods can be found, for example, in "Antibodies" (Harlow, E. and Lane, D., Antibodies. pp511-552, Cold Spring Harbor Laboratory Publications, New York (1988)). Moreover, methods based on those described above can generally be used even in the case of antibodies from other animal species.

The proteins of the present invention, which are used in the immunoprecipitation, can have a recognition site (epitope) for the monoclonal antibody, whose specificity has been well characterized, that is introduced into the N-terminus or the C-terminus of the proteins. The proteins have thus been made into fusion proteins with the epitope, and the immune complexes can be formed by reacting the antibody to the epitope.

A variety of epitope-antibody systems are commercially available, and these can also be used (Jikken Igaku (Experimental Medicine) 13:85-90 (1995)). Some commercially available vectors can express relatively large fusion proteins, such as those with β-galactosidase, maltose-binding protein, glutathione S-transferase, and Green fluorescent protein, by incorporating the DNA encoding the desired protein through multi-cloning sites. In order to minimize the changes in the properties of the desired protein due to fusing, methods have been reported in which only a small epitope portion having several to a dozen or so amino acids is inserted. For example, the epitopes in poly-histidine (His-tag), influenza hemagglutinin HA, human c-myc, FLAG, vesicular stomatitis virus glycoprotein (VSV-GP), T7 gene 10 protein (T7-tag),

human herpes simplex virus glycoprotein (HSV-tag), E-tag (an epitope on a monoclonal phage), etc., and their corresponding antibodies that recognize the epitopes can be used (Jikken Igaku 13: 85-90 (1995)). Any other epitope-antibody system can be used, as long as it can detect the fusion protein. It should be noted that the fusion proteins that bind to the proteins of the present invention can be isolated by means of affinity chromatography, without using antibodies. For example, the glutathione-Sepharose 4B column can be used for a GST-fusion protein.

SDS-PAGE is generally used to analyze the immunoprecipitated proteins. In this method, gel of an appropriate concentration is used according to the molecular weights of the proteins so that the bound proteins can be analyzed. It is generally difficult to detect the bound proteins with ordinary staining methods for proteins (e.g., the Coomassie Brilliant Blue (CBB) staining method or the silver staining method). However, the cells can be cultured in a medium to which ^{35}S-methionine or ^{35}S-cysteine has been added to label the proteins, in order to increase the detection sensitivity. Once the molecular weight of a protein becomes known, it is possible to purify the protein directly from the SDS-polyacrylamide gel and to determine its sequence. In addition to the immunoprecipitation method described above, it is also possible to prepare the proteins by running the culture supernatant or the cellular extracts of the cells expected to express the proteins that bind to the proteins of the present invention through an affinity column having the proteins of the present invention immobilized onto it, then purifying the proteins that specifically bound to the column.

It is also possible to directly screen for the genes encoding the proteins that bind by using the proteins of the present invention. In this screening method, the gene products of the sample genes are brought into contact with the proteins of the present invention, thereby selecting the genes corresponding to the gene products that bind to the proteins of the present invention. There are no restrictions on the sample genes, but cDNA libraries prepared from the cells expected to express the proteins that bind to the proteins of the present invention are preferable. A specific example of the method utilizes the yeast 2 hybrid system (Fields, S. and Song, O., Nature 340:245-247 (1989)). Namely, one can express the proteins of the present invention within the yeast cells by fusing them with the SRF binding region, GAL4 binding region, or LexA binding region. One can then introduce the cDNA libraries prepared from the cells expected to express the proteins that bind to the proteins of the present invention into the above yeast cells so that the proteins are expressed in a form fused with the VP16, GAL4 transcription activation domain, or the *E. coli* B42 peptide. Finally, one can isolate the library-derived cDNA from the positive clones. (When a protein that binds to the protein of the present invention is expressed within the yeast cell, the binding between these proteins activates the reporter gene, enabling the detection of the positive clone.)

The vectors and expression libraries to be used in this system can be purchased from several sources (Clontech, MATCHMAKER Two-Hybrid System; Stratagene, HybriZAP II Two-Hybrid System). For the specific method, one can follow the manufacturer's manual. The genes encoding the proteins that bind to the proteins of the

present invention can be obtained directly by this method. In fact, the bindings between APC and hDLG (Matsumine, A., et al., Science 272:1020-1023 (1996)), between GRIP and the AMPA receptor (Dong, H., et al., Nature 386:279-284 (1997)), between Homer and the glutamate receptor (Brakeman, P. R., et al., Nature 386:284-288 (1997)), and between SRY and SIP-1 (Poulat, F., et al. J. Biol. Chem. 272:7167-7172 (1997)) were confirmed and the target proteins of the proteins having the PDZ domain were identified using this yeast 2 hybrid system.

It is also possible to screen the proteins by the "west-western blotting method" (Skolnik, E. Y., Margolis, B., Mohammadi, M., Lowenstein, E., Fischer, R., Drepps, A., Ullrich, A., and Schlessinger, J., Cloning of PI3 kinase-associated p85 utilizing a novel method for expression/ cloning of target proteins for receptor tyrosine kinases. Cell 65:83-90 (1991)). In this method, a cDNA library is prepared using a phage vector (such as λgt11 and ZAP) from the cells expected to express the proteins that bind to the proteins of the present invention (e.g., human umbilical vascular endothelial cells). The protein is then expressed on an LB-agarose, and the expressed proteins are fixed onto a filter with which the protein of the present invention that has been biotin-labeled or purified as a fusion protein with the GST protein is reacted. The plaques expressing the binding proteins are detected with streptavidin or an anti-GST antibody. It is then also possible to introduce the isolated genes from the above procedure into *E. coli* or other cells to express them and to prepare the proteins encoded by the genes.

It should be possible to determine the signal transduction pathways mediated by the protein-protein

interaction by using the proteins of the present invention to isolate and analyze their binding proteins and the genes encoding them. Furthermore, as the relationship between the signal transduction and diseases becomes clearer, it will be possible to develop pharmaceuticals targeted at the proteins of the present invention and the proteins that interact with them.

It is also expected that treatments using antisense DNA against the DNA encoding these proteins will become possible. In the present invention, "antisense DNA" refers to the DNA encoding the RNA that is complementary to the transcription product of the target gene, thereby employing the activity to suppress the expression of the target gene. Antisense DNA does not have to be perfectly complementary to the transcription product of the target gene, as long as it can effectively block the expression of the target gene. It preferably possesses 90% or more, and more preferably 95% or more, complementarity. The chain length of the antisense DNA is 15 nucleotides or more, preferably 100 nucleotides or more, and more preferably 500 nucleotides or more. Various modified antisense oligonucleotides are being utilized as antisense DNA. For example, phosphorothioates (S-oligos) are preferable in terms of stability and solubility. The methods for introducing antisense DNA include direct administration, lipofection, the HVJ method, and the HVJ-liposome method. It is also possible to perform the treatment with antisense RNA using vectors. In this case, the gene therapy is achieved by inserting the DNA of the present invention backwards into the vector used in the recombinant protein production in animal cells described above. The DNA is

then expressed within the body by introducing it through direct administration, lipofection, the HVJ method, the HVJ-liposome method, etc. It is also possible to employ the methods of gene introduction using virus vectors such as adeno-associated virus, Adenovirus, human herpes simplex virus, vaccinia virus, and Fowlpox virus, in order to express the antisense RNA within the body. Treatments using ribozymes, instead of antisense DNA, are also possible.

Brief Description of the Drawings

Figure 1. Sequence comparisons between "32-8-1" (top) and "AF00168" (bottom) are shown.

Figure 2. Sequence comparisons between "32-8-1" (top) and "AJ001319" (bottom) are shown.

Figure 3 Sequence comparisons between "32-8-1" (top) and "AJ001320" (bottom) are shown.

Figure 4. Continuation of Figure 3, sequence comparisons between "32-8-1" (top) and "AJ001320" (bottom) are shown.

Figure 5. Photograph of the electrophoresis showing the results of northern blot analysis of the "32-8-1 gene" is shown. The BamHI-XbaI fragment was used as the probe. "H" in the figure indicates the results with the Human Multiple Tissue Northern (MTN) Blot (Clontech #7760-1). The lanes are 1. Heart, 2. Brain, 3. Placenta, 4. Lung, 5. Liver, 6. Skeletal muscle, 7. Kidney, and 8. Pancreas. "H4" indicates the results with the Human Multiple Tissue Northern (MTN) Blot IV (Clontech #7766-1), and the lanes are 1. Spleen, 2. Thymus, 3. Prostate, 4. Testis, 5. Uterus, 6. Small intestine, 7. Colon, and 8. Peripheral Blood Leukocyte. "F2" indicates the results with the Human

Fetal Multiple Tissue Northern (MTN) Blot II (Clontech #7756-1), and the lanes are 1. Fetal brain, 2. Fetal lung, 3. Fetal liver, and 4. Fetal kidney.

Figure 6. Photograph of the electrophoresis showing the results of northern blot analysis of the "32-8-1 gene" is shown. The NdeI 1.2 kb-#1 probe was used. "H" in the figure indicates the results with the Human Multiple Tissue Northern (MTN) Blot (Clontech #7760-1), and the lanes are 1. Heart, 2. Brain, 3. Placenta, 4. Lung, 5. Liver, 6. Skeletal muscle, 7. Kidney, and 8. Pancreas. "H4" indicates the results with the Human Multiple Tissue Northern (MTN) Blot IV (Clontech #7766-1), and the lanes are 1. Spleen, 2. Thymus, 3. Prostate, 4. Testis, 5. Uterus, 6. Small intestine, 7. Colon, and 8. Peripheral Blood Leukocyte. "F2" indicates the results with the Human Fetal Multiple Tissue Northern (MTN) Blot II (Clontech #7756-1), and the lanes are 1. Fetal brain, 2. Fetal lung, 3. Fetal liver, and 4. Fetal kidney. "Mu" indicates the results with the Human Muscle Multiple Tissue Northern (MTN) Blot (Clontech #7765-1), and the lanes are 1. Skeletal muscle, 2. Uterus, 3. Colon, 4. Small intestine, 5. Bladder, 6. Heart, 7. Stomach, and 8. Prostate. "C" indicates the results with the Human Cancer Cell Line Multiple Tissue Northern (MTN) Blot (Clontech #7757-1), and the lanes are 1. Promyelocytic leukemia HL-60 cells, 2. HeLa S3 cells, 3. Chronic myelogenous leukemia K-562 cells, 4. Lymphoblastic leukemia MOLT-4 cells, 5. Burkitt's lymphoma Raji cells, 6. Colorectal adenocarcinoma SW480 cells, 7. Lung carcinoma A549 cells, and 8. Melanoma G361 cells.

Figure 7. Positional relationships among various clones isolated by the present inventors are presented.

These are the "32-8-1" cDNA clone; the heart cDNA derived "686-1-2" and "686-1-4" clones. The fetal liver cDNA derived "FL #5," "#12," and "#6" clones are also shown. The PDZ domains encoded by the 32-8-1 gene are indicated by circles. The translation initiation site at nucleotide 292 and the translation termination site at nucleotide 4410 are also indicated in the figure. The positions of the probes, NdeI 1.2 kb-#1 and BamHI-XbaI, are also shown.

Figure 8. PDZ domain sequences of the protein (SEQ ID NO: 1) encoded by the 32-8-1 gene are shown. The PDZ domain sequences that exist within the protein encoded by the 32-8-1 gene are aligned.

Figure 9. Four colonies of *E. coli* transformants expressing GST-PDZ56 were picked, and the expression was compared depending on the presence or absence of the isopropyl thiogalactoside (IPTG) induction. Transformants with pGST-2TK were used as a control. The samples from each clone were analyzed on a 10% to 20% SDS-polyacrylamide gel, with even-numbered lanes before the IPTG induction and odd-numbered lanes three hours after the IPTG induction. Lanes 2 and 3 correspond to the pGST-2TK transformants, and lanes 4 through 11 correspond to clone 1 through 4 of the *E. coli* transformants expressing GST-PDZ56. Lane 1 shows molecular weight markers. The bands corresponding to the induced expression of GST-PDZ56 are indicated with an arrow.

Figure 10. The same samples used in the experiment shown in Figure 9 were analyzed by western blot. Bands (indicated with an arrow) corresponding to the induced expression of the 55 kDa protein were detected with the anti-GST antibody. The bands near 30 kDa seen in the samples three hours after the IPTG induction are

interpreted to represent the degradation of the GST-PDZ56 protein.

Figure 11. The expression of GST-PDZ14 from the *E. coli* transformants three hours after the IPTG induction was analyzed by Coomassie blue staining. Lanes 2 and 6 correspond to the samples prior to the IPTG induction; lanes 3 through 6 correspond to clones 1, 2, 3, and 4 of the *E. coli* HB101 transformants; and lanes 8 through 11 correspond to clones 1, 2, 3, and 4 of the *E. coli* JM109 transformants, showing the results of GST-PDZ14 expression after the IPTG induction (arrow). Lane 1 shows molecular weight markers.

Figure 12. The purification process of PDZ56 is shown. Coomassie blue staining was used. Lane 1 shows molecular weight markers. Lane 2 corresponds to the culture media; lane 3, to the sonicated sample; lane 4, to the fraction unbound to the glutathione-Sepharose column; lanes 5 through 7, to the washes; and lanes 8 and 9, to the PDZ56 protein not containing the GST protein portion, which has come off the glutathione-Sepharose column after digestion by thrombin. Bands at approximately 30 kDa can be clearly seen (arrow indicates PDZ56). Lane 10 corresponds to the GST protein portion bound to the glutathione-Sepharose column, which was eluted after digestion by thrombin (arrow indicates GST). Lanes 11 and 12 show the GST-PDZ56 fusion protein that was eluted without thrombin digestion in a regular elution buffer containing glutathione (arrow indicates GST-PDZ56).

Figure 13. The results of western blotting performed with anti-GST antibody using a filter onto which the same samples used in the experiment shown in Figure 12 were blotted are shown. Comparing lanes 8 and 9 with lane 10

(arrow: GST) clearly shows that the 55kDa GST-PDZ56 fusion protein (arrow indicates GST-PDZ56) shown in lanes 11 and 12 has been cleaved by thrombin to yield only PDZ56 that does not contain the GST portion. Bands in lanes 8 and 9 of Figure 12 cannot be detected by the GST antibody used in Figure 13 because they do not contain GST.

Figure 14. The purification process of PDZ14, similar to that in Figure 9, is shown. Lane 1 shows molecular weight markers. Lane 2 corresponds to the culture media; lane 3, to the sonicated sample; lane 4, to the fraction unbound from the glutathione-Sepharose column; lanes 5 through 8, to the washes; and lanes 9, 10, and 11, to the PDZ14 protein not containing the GST protein portion, which came off the glutathione-Sepharose column after digestion by thrombin. Bands at 65 kDa can be clearly seen (arrow indicates PDZ14). However, degradation products of the PDZ14 protein were also detected at 28 kDa and 37 kDa (arrows indicate 37 kDa and 28 kDa).

Figure 15. Out of the Protein Medley (Clontech), the filters blotted with 100mg each of the cell lysates from human testis (T), skeletal muscle (Sk), liver (Lv), heart (H), and brain (B) were reacted with the antisera from the rabbits immunized with peptide 32-8-1-17, PDZ14, or PDZ56 for western blotting. The filters were reacted sequentially with the 5,000-fold diluted rabbit antiserum, the 1,000-fold diluted biotin-labeled anti-rabbit Ig antibody, and the 2,500-fold diluted horseradish peroxidase (HRP)-labeled streptavidin-biotin complex (Amersham). The results of detection by chemiluminescence of the proteins that react with the rabbit antisera are shown. In the liver tissue, the present inventors were able to detect a

band at around 130 kDa. This band is expected to have been derived from the 32-8-1 protein (arrow).

Figure 16. The results of an analysis of the tissue specificity of the 32-8-1 gene expression by RT-PCR are shown. The 24-types of first strand cDNAs used were 1. brain, 2. heart, 3. kidney, 4. liver, 5. lung, 6. pancreas, 7. placenta, 8. skeletal muscle, 9. colon, 10. ovary, 11. peripheral leukocyte, 12. prostate, 13, small intestine, 14. spleen, 15. testis, 16. thymus, 17. fetal brain, 18. fetal heart, 19. fetal kidney, 20. fetal liver, 21. fetal lung, 22. fetal skeletal muscle, 23. fetal spleen, and 24. fetal thymus. Single bands at 650 bp were detected in panel A, and three bands (750 bp, 850 bp, and 950 bp) were detected in panel B.

Figure 17. The comparisons among the sequences of FH750, FH850, and FH950 are shown.

Figure 18. The continuation of Figure 17 showing the comparisons among the sequences of FH750, FH850, and FH950 is shown.

Figure 19. A photograph of an electrophoresis presenting the results of detection of the 32-8-1b protein by western blotting is shown. Lanes 1 and 2 were detected with the antisera against the 32-8-1-17 peptide, and lanes 3 and 4 were detected with the antisera against PDZ56. Cell lysates from neuroblastoma cells SH-SY5Y (lanes 1 and 3) and the NT-N cells (lanes 2 and 4), which are neurons differentiated from NT-2 by the retinoic acid stimulation, were separated on an SDS-polyacrylamide gel. Bands that are expected to correspond to the 32-8-1b protein were detected with a size of 250 kDa or more.

Figure 20. Sequence comparisons between "32-8-1b" (top) and "AF00168" (bottom) are shown.

Figure 21. Sequence comparisons between "32-8-1b" (top) and "AJ001319" (bottom) are shown.

Figure 22. Sequence comparisons between "32-8-1b" (top) and "AJ001320" (bottom) are shown.

Figure 23. The continuation of Figure 22, which presents sequence comparisons between "32-8-1b" and "AJ001320," is shown.

Figure 24. The continuation of Figure 23, which presents sequence comparisons between "32-8-1b" and "AJ001320," is shown.

Figure 25. The sequences of the PDZ domains in the protein (SEQ ID NO: 83) encoded by the 32-8-1b gene are shown. The sequences of the PDZ domains that exist within the protein encoded by the 32-8-1b gene are aligned.

Figure 26. The positional relationships among the various clones isolated by the present inventors are shown. These are the "32-8-1" cDNA clone; the heart cDNA derived "686-1-2" clone, "686-1-4" clone, and "FH950" clone; the fetal liver cDNA derived "FL#5," "#12," and "#6" clones; and the brain derived "1.2 kb #33" clone and "D-2" clone are shown. The PDZ domains encoded by the "32-8-1b" gene are indicated by rectangles.

Best Mode for Implementing the Invention

Embodiments of the present invention are exemplified below. However, the present invention shall in no way be limited by these examples.

Example 1 Cloning of genes

(1) Differential display

The human umbilical vascular endothelial cells (HUVEC) were obtained from Morinaga Biochemistry Research Institute

and cultured by using the Normal Human Vascular Endothelial Cell Culturing kit (Catalog #680051). When the cells became subconfluent, 10 ng/ml Recombinant Human Tumor Necrosis Factor-α (TNFα, Catalog #300-01A, PEPROTECH Inc.) was added, and the cells were cultured for another 24 hours. The expressed genes were compared with those from the cells without the addition of TNFα. Cells were detached from the plate with trypsin-EDTA, precipitated by centrifugation at 1,000 rpm for 5 minutes, and washed once with PBS. The total RNA was then recovered by using an RNAeasy Total RNA kit (QIAGEN). Using 0.2 μg of the recovered total RNA, the present inventors synthesized cDNA by means of the H-T11G anchor primer. The conditions were based on those given in the manual for the RNAimage kit (GenHunter). Genes were randomly amplified using the TAKARA Taq polymerase through 40 cycles of polymerase chain reaction (PCR). Each cycle consisted of 94°C for 30 seconds, 40°C for 2 minutes, and 72°C for 30 seconds, for each of the eight kinds of arbitrary primers H-AP1 through H-AP8. The reaction mixture contained α-^{32}P dATP. The products were separated on sequencing gels, and those genes whose bands were intensified by the TNFα stimulation, that is, the genes whose mRNA expression was increased as compared to the case with no stimulation, were amplified again with the same conditions. The primer DNA was then removed from the reaction mixture using a Qiaquick Spin PCR Purification kit. The nucleotide sequence information of "DDEST32" shown in SEQ ID NO: 13 was obtained by analyzing the products with a Dye Terminator Cycle Sequencing FS Ready Reaction kit (Perkin Elmer, Catalog #402122) using the same primers used for amplification.

(2) Construction of cDNA library

A cDNA library was constructed using a ZAP-cDNA synthesis kit (Stratagene). A 10 x 1st strand buffer (5 µl), 3 µl of 1st strand methyl nucleotides mix, 2 µl of linker-primer (1.4 µg/µl), 1 ml of RNase Block ribonuclease inhibitor (40 U/µl), 10 µl of TNFα-stimulated HUVEC poly A^+ mRNA (0.5 µg/µl), and 24 µl of diethyl pyrocarbonate (DEPC)-treated water were gently mixed and allowed to stand at room temperature for 10 minutes. SuperScript II reverse transcriptase (5 µl, 200 U/µl, GIBCO-BRL) was mixed with the cDNA library. The mixture was incubated at 37°C for 40 minutes then at 45°C for 70 minutes. The reaction mixture was put on ice, and 20 µl of 10 x 2nd strand buffer, 6 µl of 2nd strand nucleotide mix, 115.9 µl of sterilized distilled water, RNase H (1.5 U/µl), and 11.1 µl of DNA polymerase I (9 U/µl) were mixed into 45 µ of the reacted mixture by vortexing, and the mixture was incubated at 16°C for 150 minutes. After the reaction, 23 µl of blunting dNTP mix and 2 µl of cloned Pfu DNA polymerase (2.5 U/µl) were added, and the mixture was incubated at 72°C for 30 minutes. The mixture was then sequentially extracted with 200 µl of phenol/chloroform, and with chloroform, and further precipitated by adding 20 µl of 3M sodium acetate and 400 µl of 100% ethanol. After overnight incubation at -20°C, the mixture was centrifuged at 15,000 rpm for 60 minutes (4°C), and the precipitate was washed with 500 µl of 70% ethanol and dried. The precipitate was dissolved in 9 µl of 0.4 µg/µl EcoRI adapter and incubated at 4°C for 45 minutes. 10 x ligase buffer (1 µl), 1 µl of 10 mM ATP, and 1 µl of T4 DNA ligase (4 U/µl) were then added to the above, and the ligation reaction was performed overnight at 8°C. The mixture was incubated at 70°C for

30 minutes to inactivate the enzyme, spun down to collect the solution on the bottom of the tube, and let sit for 5 minutes at room temperature. To the mixture were added 1 µl of 10 x ligase buffer, 2 µl of ATP, 6 µl of sterilized water, and 1 µl of T4 polynucleotide kinase (10 U/µl). The mixture was incubated at 37°C for 30 minutes then incubated again at 70°C for 30 minutes to inactivate the enzyme. XhoI buffer supplement (28 µl) and 3 µl of XhoI (40 U/µl) were added to the above, and the mixture was reacted at 37°C for 90 minutes. The mixture was cooled to room temperature, then 5 µl of 10 x STE buffer was added. The mixture was then applied to a Sephacryl S-500 column and eluted twice with 60 µl of 1 x STE buffer. Ethanol (120 ml) was then added to the mixture, and the mixture was allowed to stand at -20°C overnight. It was then centrifuged at 15,000 rpm for 60 minutes (4°C) to obtain the precipitate. The precipitate was washed with 200 µl of 80% ethanol and dried. It was then dissolved with 6 µl of sterilized water, and 2.5 µl of it was used for the ligation reaction with the vector. To 2.5 µl of the cDNA, 1 µl of Uni-ZAP XR vector (1 µg), 0.5 µl of 10 x ligase buffer, 0.5 µl of 10 mM ATP, and 0.5 µl of T4 DNA ligase (4 U/µl) were added and reacted at 12°C overnight. The ligation mixture (1 µl) was added to the GigapackIII Gold Packaging extract, mixed well, and incubated for two hours at room temperature. SM buffer (500 µl; 5.8 g NaCl, 2.0 g $MgSO_4$-$7H_2O$, 50 ml 1 M Tris-HCl (pH 7.5), and 5 ml 2% (w/v) gelatin, brought up to 1 L with deionized water) was added to the above, and after 20 µl of chloroform was added, it was gently mixed. The mixture was then centrifuged, and the supernatant was transferred to another tube and stored at 4°C. The phage titer was measured using 0.1 µl and 1 µl

of the packaging reaction. Since approximately 300 plaques were obtained from 0.1 µl, the titer was estimated to be 3,000 PFU (plaque-forming units) per microliter. XL1 Blue MRF' was used as the host *E. coli*. It was cultured in 20 ml LB/ 10 mM $MgSO_4$/ 0.2% maltose at 37°C for 6 hours, placed on ice for 5 minutes before the OD_{600} became 1.0, and centrifuged at 500 x g for 10 minutes. To resuspend the precipitated cells, 10 ml of 10 mM $MgSO_4$ was added, and cells were diluted with 10 mM $MgSO_4$ so that the OD_{600} became 0.5. The packaging reaction (17 µl) was added to 600 µl of the freshly prepared XL-1 Blue MRF', and incubated at 37°C for 15 minutes. NZY top agar (6.5 ml; made by adding 0.7% (w/v) agarose to the NZY medium and autoclaving it), which had been incubated at 45°C, was added to the above and plated onto NZY agar plates. The plates were prepared as follows. NaCl (5 g), $MgSO_4$-$7H_2O$ (2.0 g), yeast extracts (5 g), NZ amines (10 g), and agar (15 g) were dissolved in deionized water to make the total volume 1 L. The solution pH was adjusted with NaOH to 7.5, after which the solution was autoclaved and poured into sterilized culture plates. After culturing at 37°C for six hours, the plaques were transferred onto a Hybond N^+ filter (Amersham, RPN203B) by placing the filter on the plate, denatured with 1.5 M NaCl-0.5 M NaOH for 7 minutes, neutralized by treating with 1.5 M NaCl-0.5 M Tris-HCl (pH 7.2)/ 1 mM EDTA for 5 minutes, and finally rinsed with 2 x SSC. After the filter was dried, the plaques were fixed onto the filter by StrataLinker (Stratagene).

(3) Screening of the cDNA library

The "DDEST32" DNA fragment was isolated on a 2% agarose gel, and recovered from an agarose gel slice with a

QIAEX II gel extraction kit (QIAGEN). The fragment was labeled by random labeling so it could be used as the probe. Using a Megaprime kit (Amersham, RPN1607), 5 µl of primer solution was added to 25 ng of the probe DNA, and incubated at 95°C for 5 minutes. After the solution was incubated at room temperature, 10 µl of labeling buffer, 18 µl of water, 5 µl of α^{32}P dCTP (-3000ci/mmol; Amersham), and 2 µl of Klenow fragment were added to it, and the mixed solution was incubated at 37°C for 30 minutes. The reaction was stopped by adding 2 µl of 0.5 M EDTA, and the free α-^{32}P dCTP was removed with a Pharmacia ProbeQuant G-50 column. After prehybridization at 60°C in the Rapid hybri buffer (Amersham, RPN1636), the labeled probe was heat-denatured at 95°C, rapidly chilled on ice, and added to the hybridization buffer. Hybridization was then performed by shaking at 60°C for two hours. The probe was used at a concentration of 2 x 10^6 cpm/ml. The filter was washed three times in 2 x SSC/ 0.05% SDS at room temperature for ten minutes each, and twice more in 0.1 x SSC/ 0.1% SDS at 60°C for 20 minutes each. The phage collected from positive plaques was diluted in SM buffer and plated onto 10-cm plates so that approximately 100 plaques were formed per plate. The secondary and tertiary screenings were performed similarly. As a result, clone "#32-8-1" was obtained as the positive clone. The gene that had been cloned in the Uni-ZAP vector was recovered as ordinary plasmid DNA by the *in vivo* excision method.

Example 2 Sequence determination of the "32-8-1" gene

(1) Preparation of the cDNA library for RACE

The cDNA for RACE was synthesized using a Marathon cDNA amplification kit (Clontech). The total RNA (1 µg)

obtained from the TNFα-stimulated HUVEC cells was used for the experiment. Oligo dT primer (1 µl, 10 µM) was added to the above, the total volume was brought to 5 µl, and the mixture was incubated at 70°C for 2 minutes and then placed on ice for 2 minutes. Two microliters of 5 x 1st strand buffer, 1 µl of 10 mM dNTP mix, and 1 µl of 100 U/µl MMLV reverse transcriptase were added to the above, and the total volume was made 10 µl. The mixture was then incubated at 42°C for 1 hour to synthesize the first strand cDNA. Sixteen microliters of 5 x 2nd strand buffer, 1.6 µl of 10 mM dNTP mix, and 4 µl of 20 x 2nd strand enzyme cocktail were next added to the mixture, the total volume was adjusted to 80 µl with water, and the mixture was incubated at 16°C for 90 minutes. T4 DNA polymerase (2 µl, 5 U/µl) was then added, and the reaction was performed at 16°C for 45 minutes. After 4 µl of 20 x EDTA/ glycogen was added to the mixture, it was deproteinized with equal volumes of phenol/chloroform, and isoamyl alcohol/chloroform.

Ethanol precipitation was done with 35 µl of 4 M ammonium acetate and 263 µl of 95% ethanol, and the precipitate was washed with 80% ethanol and spontaneously dried for 10 minutes. The dried precipitate was dissolved in 10 µl of deionized water, and 7.5 µl was used for the adapter ligation reaction. Marathon cDNA adapter (3 µl, 10 µM), 3 µl of 5 x DNA ligation buffer, and 1.5 µl of T4 DNA ligase (1 U/µl) were added to the above and reacted overnight at 16°C. The enzyme was inactivated by incubation at 70°C for 5 minutes, and the total volume was adjusted to 150 µl by using 135 µl of Tricine-EDTA buffer contained in the kit.

(2) cDNA cloning by RACE and sequence determination

Clone #32-8-1 was subcloned by utilizing the restriction enzyme recognition sites within the gene (PstI, XbaI, BamHI, and EcoRI), and the nucleotide sequence was determined by the cycle sequence method using a Dye Terminator Cycle Sequencing FS Ready Reaction kit (Perkin Elmer, Catalog #402122).

The sequences of the primers used are shown in Table 1. "C" indicates a primer for the complementary DNA strand.

Table 1

Primer #		DNA sequence	SEQ ID NO:
106	C	CTCCCCATCCCTCGTCCACC	14
XE	C	CTCTGACTCTGACTGACTGG	15
EX		ATGAGTTTGGTTACAGCTGG	16
402		TCAGAGAGCGTTATGGAACC	17
XER		AGTCTTGCTGGGAACAAAGA	18
801		ACTGTTACTACTTCTGATGC	19
1192-1161		TCTGATGGTCCCACAGTCTG	20
1282	C	GTTGTTTCGCAGCCAGGGAT	21
1524		CTGAGCATCGTTGGGGGTTC	22
1449	C	CCTCATCTCTGTAGAGTGTC	23
1683		TGTTAGCCCCCTCACTAAGG	24
1803		GCTATGTGCTAGGAAATACG	25
2116		TAGGGAGAAGGATCAGAGCG	26
607-93		ACAGATTTCTGACTCACTGG	27
128		TGGAAATAGGCATTCTTCAG	28
607-462		ATACAAAGACGGTCTAATCC	29
2920	C	CCGCTTTCCCATCTTTAGAAAC	30
3121		TATCTCGTGTGGAAGATGTG	31
2266-107	C	ACATAAATGTTGCTATCACC	32
3361		TGCCACTTAGTAGCCGAGTG	33
3615		GCATTGCATTACAGTTGAGC	34
1301	C	TCCTCCTTTGACAATGTCTG	35
BXR	C	CATTTCGACTGTTCTTAATC	36
XB	C	TCAGTGGATGTGCCACAGAT	37
4221	C	CAGTAGGTTAACTGCTTCGG	38
BX		AGTTCCAGTCTTTCTTTCGG	39
4335		TTTCTTTCACTGGGCTGAAGTC	40
XBR		CCTCTGAAGACGGACGTCTG	41

Accordingly, a nucleotide sequence of 5,146 bp was determined. When the first G of the EcoRI recognition site was counted as nucleotide 1, the PDZ domain started at nucleotide 468. A stop codon immediately followed three repeated stretches of approximately 80 amino acids. The sequence in the 3′ region of the gene also contained three PDZ domains at a distance of approximately 2 kb from the stop codon described above. (An experiment conducted later revealed that clone #32-8-1 contains a sequence of approximately 2 kb. This sequence has been derived from an intron, transcribed and inserted, thereby introducing the stop codon immediately after the first three PDZ domains.)

Therefore, the present inventors performed 5′ Rapid Amplification of cDNA Ends (RACE) starting from the position of the three PDZ domains found in the latter half. Using 5 µl of the cDNA described above, 5′ RACE was performed according to the manual contained in the kit. The reaction mixture consisted of 5 µl of cDNA, 5 µl of 10 x Advantage™ KlenTaq buffer (which came with the kit), 4 µl of 2.5 mM dNTP, 1 µl of 10 µM AP1 primer (CCATCCTAATACGACTCACTATAGGGC (SEQ ID NO: 42)), 1 µl of 10 µM 32-8-1 5′ RACE primer #22(TTGGGGTGGGGAGAGGAGGTAGATTGC (SEQ ID NO: 43)), 1 µl of Advantage™ KlenTaq polymerase mix (Toyobo, CLK8417-1), and 33 µl of deionized water to make the total 50 µl. PCR reaction was performed using a Perkin Elmer Thermal Cycler 2400. A reaction consisting of 94°C for 1 minute, five cycles of 94°C for 5 seconds and 72°C for 2 minutes, five cycles of 94°C for 5 seconds and 70°C for 2 minutes, followed by 25 cycles of 94°C for 5 seconds and 68°C for 2 minutes did not produce clearly detectable bands. By performing nested PCR under the same conditions, the present inventors were able to obtain a band of

approximately 1.8 kb. Here the AP2 primer (ACTCACTATAGGGCTCGAGCGGC (SEQ ID NO: 44)) and 32-8-1 5' RACE primer #1034 (GCACATCACCAAGTGGGCTGCCTACTC (SEQ ID NO: 45)) were used as primers, and 5 µl of the 50-fold dilution of the initial PCR product was used. Also, the original 25 cycles of 94°C for 5 seconds and 68°C for 2 minutes was reduced to 15 cycles. As a result, cDNA clone "32-8-1/5R3," which does not contain the 2 kb gap, was obtained.

Next, the present inventors determined the sequence of clone 32-8-1/5R3. The sequences of the primers used for the sequence determination of 32-8-1/5R3 are shown in Table 2. "C" indicates a primer for the complementary DNA strand.

Table 2

Primer #		DNA sequence	SEQ ID NO:
EX		ATGAGTTTGGTTACAGCTGG	46
456	C	AATCTAATGCAGCTCGCCTG	47
XER		AGTCTTGCTGGGAACAAAGA	48
678	C	TCACTTTAGAAGGGGCACAT	49
801		ACTGTTACTACTTCTGATGC	50
1192-1161		TCTGATGGTCCCACAGTCTG	51
1282	C	GTTGTTTCGCAGCCAGGGAT	52
1524		CTGAGCATCGTTGGGGGTTC	53
1449	C	CCTCATCTCTGTAGAGTGTC	54
2116		TAGGGAGAAGGATCAGAGCG	55
1301	C	TCCTCCTTTGACAATGTCTG	56
839		TTTCATCATCTACAGCCAGT	57
1389		TGACACCCTCACTATTGAGC	58

The nucleotide sequence of 2,819 bp, which was determined by combining the sequences of clones #32-8-1 and 32-8-1/5R3, is shown in SEQ ID NO: 59.

Example 3 Cloning of a cDNA clone corresponding to the 5' upstream region of the 32-8-1/5R3 cDNA clone by RACE

The present inventors attempted to isolate the upstream cDNA located 5' to the 32-8-1/5R3 clone by the 5' Rapid Amplification of cDNA Ends (RACE) method. A human heart cDNA library and a human fetal liver cDNA library were used as cDNA sources. Two clones, 2.8 kb and 1.2 kb in size, were obtained from the human heart cDNA library. One 1.1 kb clone was obtained from the human fetal liver cDNA library. The cloning procedure is described below.

The present inventors used cDNA Library Human Heart (Takara Shuzo, Catalog #9604) for the human heart cDNA library. The XL1 Blue-MRF' *E. coli* cells transformed with the plasmid DNA containing the cDNA inserted into the pAP3neo vector (Genbank Accession No.AB003468) were cultured by the usual method, the plasmid DNA was recovered by the alkaline method, and the cDNA clone containing the 5' upstream region was obtained by PCR using 10 ng of the plasmid DNA as the template. The reaction mixture consisted of 10 ng of the cDNA, 5 µl of 10 x Advantage™ KlenTaq buffer (which came with the kit), 4 µl of 2.5 mM dNTP, 1 µl of 10 µM AP3neo5' primer (which came with the kit; 5'-GCCCTTAGGACGCGTAATACGACTC-3' (SEQ ID NO: 60)), 1 µl of 10 µM 32-8-1 5' RACE primer #686 (5'-AGCCAGTATCTGATCTCCGACTTTG-3' (SEQ ID NO: 61)), and 1 µl of Advantage™ KlenTaq polymerase mix (Toyobo, CLK8417-1). These were mixed with deionized water to make the total 50 µl. PCR reaction was performed using a Perkin Elmer Thermal Cycler 2400. A reaction consisting of 94°C for 1 minute, 5 cycles of 94°C for 5 seconds and 72°C for 4 minutes, and 5 cycles of 94°C for 5 seconds and 70°C for 4 minutes, followed by 25 cycles of 94°C for 5 seconds and

68°C for 4 minutes, yielded the bands of 2.8 kb and 1.2 kb. The products were separated on a 0.8% agarose gel. The corresponding bands were excised and purified with the QIAquick gel extraction kit (QIAGEN, 28706) and were subjected to TA cloning according to the manual for the pGEM-T Vector System I (Promega, A3600). The present inventors designated the two clones as 686-1-4 (2.8 kb) and 686-1-2 (1.2 kb). The sequence of clone 686-1-2 is contained in that of 686-1-4 (sequence 686-1-4), and ranges from nucleotide 1585 to nucleotide 2793 of SEQ ID NO: 3 (Figure 7).

The present inventors performed 5' RACE using Marathon Ready human fetal liver cDNA (Clontech) as the human fetal liver cDNA library. The reaction mixture consisted of 5 µl of the cDNA, 5 µl of 10 x Advantage™ KlenTaq buffer (which came with the kit), 4 µl of 2.5 mM dNTP, 1 µl of 10 µM AP1 primer (which came with the kit: 5'-CCATCCTAATACGACTCACTATAGGGC-3' (SEQ ID NO: 42)), 1 µl of 10 µM 32-8-1 5' RACE primer #686 (5'-AGCCAGTATCTGATCTCCGACTTTG-3' (SEQ ID NO: 61)), and 1 ml of Advantage™ KlenTaq polymerase mix (Toyobo, CLK8417-1). These were mixed with 33 µl of deionized water to make the total 50 µl. PCR reaction was performed using a Perkin Elmer Thermal Cycler 2400. A reaction consisting of 94°C for 1 minute, 5 cycles of 94°C for 5 seconds and 72°C for 6 minutes, 5 cycles of 94°C for 5 seconds and 70°C for 6 minutes, and 25 cycles of 94°C for 5 seconds and 68°C for 6 minutes did not produce clearly detectable bands. The reaction mixture was then diluted 50 fold, 5 µml of which was mixed with 5 µml of 10 x Advantage™ KlenTaq buffer (which came with the kit), 4 µl of 2.5 mM dNTP, 1 µl of 10 µM AP2 primer (which came with the kit;

5'- ACTCACTATAGGGCTCGAGCGGC-3 (SEQ ID NO: 44)), 1 µl of 10 µM 32-8-1 5' RACE nested primer #FLN (5'-ATTTTCACTTTAGAAGGGGCACAT-3'(SEQ ID NO: 62)), 1 µl of Advantage™ KlenTaq polymerase mix (Toyobo, CLK8417-1), and 33 µl of deionized water to make the total 50 µl. Nested PCR was performed at 94°C for 1 minute, 5 cycles of 94°C for 5 seconds and 72°C for 6 minutes, 5 cycles of 94°C for 5 seconds and 70°C for 6 minutes, and 15 cycles of 94°C for 5 seconds and 68°C for 6 minutes, which produced a band of approximately 1.1 kb. The products were separated on a 0.8% agarose gel. The corresponding bands were then excised and purified with a QIAquick gel extraction kit (QIAGEN, 28706) and were subjected to TA cloning according to the manual for the pGEM-T Vector System I (Promega, A3600). Three different clones were thus obtained and designated HFL#5, HFL#12, and HFL#6. HFL#5 and HFL#12 started from nucleotide 1357 of SEQ ID NO: 3, while HFL#6 started from nucleotide 1377 of SEQ ID NO: 3. Of course, all three contained the sequence up to that of primer #FLN, which was used in the RACE (Figure 7).

The nucleotide sequences were determined as described above, by means of the cycle sequence method using a Dye Terminator Cycle Sequencing FS Ready Reaction kit (Perkin Elmer, Catalog #402122). The combined sequence of the previously determined one and the newly determined one is shown in SEQ ID NO: 3. Figure 8 shows the sequences of nine PDZ domains aligned. The primers used for the cycle sequencing method are listed in Table 3.

Table 3

Primer #	DNA sequence	SEQ ID NO:
686A	GGCATAACTTTACTTACTTG	63
686B	ATCTACTAAGTCAGCATCAT	64
686C	ATTTGCAGGTGTGTAGTCAT	65
686D	TTCCTTCTGTGCTACCCGAT	66
686E	GGACTATCTTCCAGAACATG	67

Example 4 Search for proteins having homology to the protein encoded by the "38-2-1" gene

The BLASTN search and the BLASTP search detected "Mus musculus 90RF binding protein 1 (9BP-1) mRNA, partial cds." (LOCUS: MMAF000168, ACCESSION: AF000168) that consists of 2703 bp as a gene having homology to the "32-8-1" gene. This gene was recorded toGenBank on 18 MAY 1997. The amino acid sequence of the protein encoded by the "32-8-1" gene (the sequence after amino acid 847 of SEQ ID NO: 1) and that of AF000168 are aligned and shown in Figure 1. In the figure, the amino acid 847 of SEQ ID NO: 1 was regarded as the "first" amino acid, and comparisons are shown with the amino acid sequence thereafter.

In addition, "Rattus norvegius mRNA for multi PDZ domain protein" (LOCUS: RNMUPP1, ACCESSION: AJ001320) consisting of 7516 bp, and "Homo sapiens mRNA for multi PDZ domain protein" (LOCUS: HSMUPP1, ACCESSION: AJ001319) consisting of 1768 bp were detected as genes having homology. These genes were registered on 26 MAR 1998. The amino acid sequence of the protein encoded by the "32-8-1" gene (the sequence after amino acid 921 of SEQ ID NO: 1) and that of AJ001319 are aligned and shown in Figure 2. The amino acid sequence of the protein encoded

by the "32-8-1" gene (the sequence of SEQ ID NO: 1) and that of AJ001320 are aligned and shown in Figures 3 and 4.

Example 5 Analysis of tissue specificity of expression by northern blotting

Clontech Human Multiple Tissue Northern (MTN) Blot (Catalog #7760-1), Human MTN Blot IV(Catalog #7766-1), Human Fetal MTN Blot II (#7756-1), Human Muscle MTN Blot (#7765-1) and Human Cancer Cell Line MTN Blot (#7757-1) were used to analyze the tissue specificity of gene expression. Northern blot was performed according to the standard method, using the BamHI-XbaI fragment (from position 3709 to position 4337 of SEQ ID NO: 3) as the probe (see Figure 7 for the position of the probe), and 25 ng of the DNA fragment was labeled with α-^{32}P dCTP using a Megaprime DNA labeling kit (Amersham, Catalog RPN1607). These MTN Blots were prehybridized in 5 ml of the ExpressHyb hybridization solution (Clontech, Catalog #8015-2) at 68°C for 30 minutes, and then hybridized with 1 x 10^7 cpm of the labeled probe also in 5 ml of the ExpressHyb hybridization solution (2 x 10^6 cpm/ ml) at 68°C for 2 hours. The filters were washed three times in 2 x SSC (0.3 M NaCl, 0.03 M sodium citrate (pH 7.0))/ 0.05% SDS at room temperature for 10 minutes each, washed twice more in 0.1 x SSC/ 0.1% SDS at 50°C for 15 minutes each, exposed on FUJI imaging plates overnight, and analyzed by a FUJI BAS2000. As shown in Figure 5, strong expression of the approximately 8 kb transcription product was detected in the heart, placenta, skeletal muscle, fetal brain, fetal lung, fetal kidney, small intestine, bladder, stomach, prostate, HeLa S3 cells, lung cancer A549 cells, and melanoma G361 cells. However, the expression was either

absent or weak in the lung, lymphoid tissues (spleen and thymus), and cell lines (lanes 1, 3, 4, 5, and 6 of blot C). In the heart, liver, kidney, and fetal liver, a 5.5 kb transcription product was expressed.

Similarly, northern blot analysis was performed using the NdeI 1.2 kb-#1 probe (from position 1 to position 1091 of SEQ ID NO: 3) (see Figure 7 for the position of the probe). However, the band corresponding to the 5.5 kb transcription product was not detected (Figure 6). Considering this and the fact that the cDNAs cloned by 5′ RACE from the fetal liver only contained the 5′ sequences 1,357 and 1,377 nucleotides downstream from the 5′ end of the transcription product expressed in the heart (Figure 7), it can be deduced that the difference of the transcription initiation sites in the heart and liver caused the difference in the lengths of the transcription products. Therefore, the peptide encoded by the 32-8-1 gene that is expressed in the liver is expected to start with the first methionine encoded by the ATG codon beginning with the 1396th nucleotide. This results in the transcription product from the liver consisting of 1,005 amino acids, compared with that from the heart consisting of 1,373 amino acids. Consequently, it does not contain PDZ domain E and is shorter by 368 amino acids. Although the biological significance of not having PDZ domain E is unclear at present, it is highly possible that this protein, by lacking this portion, is involved in a different signal regulation in the liver cells than in the other tissues since PDZ domains are important for protein-protein interactions

Example 6 Expression of the 32-8-1 protein in *E. coli*

(1) Construction of the expression vector

In order to express the 32-8-1 protein in *E. coli* as a fusion protein with the glutathione-S-transferase (GST) protein, part of the 32-8-1 gene was ligated to the carboxyl terminus of the GST gene in Pharmacia's pGEX-2TK (Genbank Accession U13851). The vector was constructed based on the method of W. Dietmaier et al. for the di-/tri-nucleotide sticky end cloning described in the PCR Application Manual (Boehringer Mannheim). pGEX-2TK (1 µg) was reacted in a mixture of 2 ml of 10 x High buffer and 20 units each of restriction enzymes EcoRI and BamHI in a total volume of 20 µl at 37°C for 3 hours. Proteins were removed by using a QIAquick column (QIAGEN) according to the manual, and the purified DNA was eluted with 30 µl of distilled water. 10 x Klenow buffer (3 µl; 100 mM Tris-HCl (pH 7.5), 70 mM $MgCl_2$, 1 mM DTT), which comes with Takara's Klenow enzyme, and 1.5 µl of 2 mM dGTP were mixed with 27 µl of the above. Four units of Klenow enzyme were then added, and the reaction was allowed to proceed at room temperature for 15 minutes. After the enzyme was inactivated by heating at 75°C for 15 minutes, the DNA was purified by deproteinizing with a QIAquick column (QIAGEN) according to the manual.

The region of the 32-8-1 gene to be expressed, which encodes amino acids 1112 to 1373, was amplified by PCR using 50 ng of #32-8-1 DNA as the template. The amplification reaction was done by adding 5 µl of 10 x Reaction buffer #1 for KOD DNA polymerase (Toyobo), 5 µl each of 10 µM primer 502-508 (5'-ATCGGGTCCATTCCATTCAGAGAGG-3' (SEQ ID NO: 68)) and 10 µM primer 758-763E

(5'-AATTGTCAAGAGAGAACCATCAAAGTGG-3'(SEQ ID NO: 69)), 4 µl of 2.5 mM dNTP, 2 µl of 25 mM $MgCl_2$, and 27 µl of sterilized water. 2.5 µl of KOD DNA polymerase was then mixed into the solution, which was then incubated at 94°C for 2 minutes and subjected to 25 cycles of 98°C for 15 seconds, 65°C for 2 seconds, and 74°C for 30 seconds. Using the QIAquick PCR purification kit, the 798 bp PCR product was purified according to the manual. The purified PCR fragment (2 µl) was mixed with 7 µl of the Boehringer's 5 x T4 DNA polymerase buffer (330 mM Tris-acetate, pH 8.0; 660 mM potassium acetate; 100 mM magnesium acetate; and 5 mM DTT), 1.5 µl of 2 mM dCTP, and 21.5 µl of sterilized water. Three units of T4 DNA polymerase were then added, and the solution was reacted at 12°C for 30 minutes. The present inventors inactivated the enzyme by incubation at 80°C for 15 minutes and purified the mixture with the QIAquick PCR purification kit according to the manual. The pGEX-2TK plasmid, which had been digested with restriction enzymes EcoRI and BamHI and modified with Klenow, and the T4 polymerase-treated PCR product were ligated by one unit of T4 DNA ligase (Promega) using the attached buffer (30 mM Tris-HCl (pH 7.8), 10 mM $MgCl_2$, 10 mM DTT, and 1 mM ATP) at 15°C overnight. The reaction mixture was then used to transform *E. coli* DH5 alpha. The recombinant protein expressed by the transformant was designated GST-PDZ56.

Similarly, a PCR product encoding the amino acids from position 611 to position 1142 of SEQ ID NO: 1 was prepared by the above described method with primer 1-7 (5'-ATCGATGGGTAGTAATCACACACAG-3' (SEQ ID NO: 70)) and primer 527-532E (5'-AATTGCTATACTGGATCCAGAGAGTGG-3' (SEQ ID NO: 71)), using clone 32-8-1/5R3 as the template. The preparation was treated with T4 polymerase under the same

conditions as before, purified with the EcoRI and BamHI-digested and Klenow-treated pGEX-2TK, and ligated. The reaction mixture was used to transform *E. coli* DH5 alpha. The recombinant protein expressed by the transformant was designated GST-PDZ14.

The *E. coli* transformants that express GST-PDZ56 were selected by the following method. Four colonies of the *E. coli* transformants obtained from the above were shake-cultured at 37°C overnight in 2 ml of the LB medium (5 g of Bacto-yeast extract (DIFCO), 10 g of Bacto-trypton (DIFCO), and 10 g of NaCl made to 1 L by dissolving them in distilled water) containing 100 µg/ml ampicillin. The solution was then diluted 100 fold in the medium of the same composition, and isopropyl thiogalactoside (IPTG) was added to a final concentration of 0.1 mM. The solution was then shake-cultured at 37°C for 3 hours. A 100 µl sample was precipitated by centrifugation at 15,000 rpm for 10 seconds and analyzed on a 10% to 20% SDS-polyacrylamide gel. The subsequent Coomassie staining easily detected the IPTG-induced expression of the approximately 55 kDa GST fusion protein from every transformant (Figure 9). Furthermore, western blotting with an anti-GST antibody also confirmed the induced expression of the 55 kDa protein band (Figure 10). For detection, the proteins in the samples separated on the 10% to 20% SDS-polyacrylamide gel were transferred onto Immobilon-P (Millipore) using a Semidry blotter (Bio-Rad) according to the methods described in the manual. The filter was then blocked at 4°C overnight with 5% skim milk (DIFCO), 2.5% bovine serum albumin (Sigma, A5940), and T-TBS (20 mM Tris-HCl (pH 7.5), 150 mM NaCl, and 0.05% Tween 20). It was next reacted at room temperature for 1 hour with the anti-sheep GST

antibody (Pharmacia) diluted 1,000 fold in the antibody dilution buffer (1% skim milk, 0.5% bovine serum albumin, T-TBS), then reacted at room temperature for 1 hour with the alkaline phosphatase-labeled anti-sheep IgG antibody diluted 1,000 fold in the antibody dilution buffer. Finally, the protein was detected by a GST Detection Module (Pharmacia).

The GST-PDZ14 was similarly expressed. However, *E. coli* HB101 and JM109 were used as hosts because *E.coli* DH5 did not produce an efficient IPTG-induced expression. The results shown in Figure 11 indicate that *E. coli* HB101 did not produce very large amounts of expression products, but that the GST-PDZ14-derived band near 90 kDa was highly induced in *E.coli* JM109. *E. coli* JM109 was subsequently used to express and purify the fusion protein.

(2) Expression and purification of the GST 32-8-1 fusion protein

The present inventors followed the method for preparing fusion proteins described on page 217 of the Supplement to Jikken Igaku (Experimental Medicine) "Shin Idenshi Kogaku Handbook (New Genetic Engineering Handbook) edited by Masami Muramatsu, et. al." published by Yohdosha to express and purify the GST fusion proteins. GST-PDZ14 and GST-PDZ56 were each cultured at 37°C for 1 hour in 2 L of LB medium to which IPTG was added to achieve a final concentration of 0.1 mM, then shake-cultured at 25°C for 5 hours. The cells were collected at 7,000 rpm for 10 minutes, resuspended in a sonication buffer consisting of PBS and 1% Triton X-100, and sonicated 1 minute for five times while chilling. The supernatant was obtained by centrifugation at 10,000 rpm for 15 minutes. It was then

applied onto a glutathione-Sepharose column, washed well with PBS, and purified with the GST Purification Module elution buffer (Pharmacia Biotech).

The 32-8-1 gene was inserted into the multi-cloning site of the pGEX-2TK expression vector (Pharmacia). This vector has a region encording the string of amino acids "Leu-Val-Pro-Arg-Gly-Ser" recognized by thrombin protease in frame after GST protein gene, so the GST protein portion can be separated from GST-32-8-1 fusion protein by applying the thrombin protease that recognizes this sequence and digesting the protein. This is useful for preparing antibodies against the protein encoded by the 32-8-1 gene (the 32-8-1 protein). Because the glutathione-Sepharose column binds to the GST protein, it was possible to purify the PDZ14 and PDZ56 portions only as the fractions not binding to the glutathione-Sepharose (Figures 12, 13, and 14) by applying the protein solution that had been digested with the thrombin protease onto the glutathione-Sepharose column. The 55 kDa GST-PDZ56 protein bands seen in Figure 12 (lanes 11 and 12) were digested by thrombin into the 25 kDa GST protein and the 30 kDa PDZ56 protein (lane 10). Furthermore, the results of western blotting using the anti-GST antibody indicated that the anti-GST antibody reacted only with the 55 kDa and 25 kDa proteins, both of which contained the GST protein (Figure 13). Together, these confirmed that the PDZ56 protein portion was cut off as the 30 kDa band (lanes 8 and 9). Similarly to GST-PDZ14, the 90 kDa GST-PDZ14 can be separated by thrombin digestion into the 25 kDa GST protein and the 65 kDa PDZ14 protein portion, as shown in Figure 14. Therefore, the protein was purified according to the following procedure. The present inventors used the method

described in the Supplement to Jikken Igaku (Experimental Medicine) "Shin Idenshi Kogaku Handbook (New Genetic Engineering Handbook) edited by Masami Muramatsu, et. al" published by Yohdosha to culture *E. coli* cells. They then used the supernatant of the sonicated cells to digest the proteins by thrombin. The detailed method followed the procedure described under Thrombin Cleavage on page 16 of the GST Gene Fusion System (Pharmacia). They next added 10 µl (10 cleavage units) of thrombin per 1 mg of the fusion protein and incubated the mixture at room temperature for 16 hours to separate the PDZ14 protein or the PDZ56 protein portion from the GST portion. By letting the cleaved GST protein portion bind to the glutathione-Sepharose column (Pharmacia), they recovered 0.56 mg of the PDZ14 or 3.5 mg of the PDZ56 protein portion as the flow-through from the column.

(3) Preparation of polyclonal antibodies using the antigens expressed in *E. coli*

The present inventors obtained polyclonal antibodies by immunizing two rabbits each with the purified PDZ14 or PDZ56 antigen. The initial immunization was done by subcutaneously injecting 0.5 mg of PDZ56 or 0.22 mg of PDZ14 bound with the carrier protein per animal as an antigen emulsion mixed with an equal amount of Freund's complete adjuvant (FCA) by the standard method. Booster injections of 0.25 mg of PDZ56 or PDZ14 were given subcutaneously as an antigen emulsion mixed with an equal amount of Freund's incomplete adjuvant (FICA) three times at two-week intervals. The proteins used as antigens were separated by SDS-PAGE and transferred onto a PVDF membrane

(Immobilon-P, Millipore); the reactivity was confirmed by western blotting.

(4) Preparation of polyclonal antibodies using a peptide

Iwaki Glass synthesized Peptide 32-8-1-17 (SEQ ID NO: 72) consisting of 21 amino acids under contract. The Keyhole limpet hemocyanin (KLH) protein was coupled to the peptide as a carrier protein by the Sulfo-MBS method, and two rabbits were immunized with the product. The initial immunization was done by subcutaneously injecting 0.4 mg of the 32-8-1-17 peptide bound with the carrier protein per animal as an antigen emulsion mixed with an equal amount of Freund's complete adjuvant (FCA) by the standard method. The second through the fifth immunizations were given at two-week intervals by subcutaneously injecting 0.2 mg of the 32-8-1-17 peptide bound with the carrier protein as an antigen emulsion mixed with an equal amount of Freund's incomplete adjuvant (FICA). The antibody titers were measured using an ELISA plate coated with the 32-8-1-17 peptide; the antisera were obtained when the titer had risen sufficiently.

(5) Reactivity of the polyclonal antibodies

Antisera were obtained by immunizing the rabbits with peptide 32-8-1-17 or with PDZ14 or PDZ56 protein expressed as a GST-fusion protein and then digested with thrombin to retain only the 32-8-1 gene product. The reactivity of the antisera was detected by western blotting using Protein Medley manufactured by Clontech. More specifically, 100 µg each of the cell lysates of the tissues from the human Testis (T), Skeletal Muscle (Sk), Liver (Lv), Heart (H), and Brain (B) of the Protein Medley manufactured by

Clontech was separated on a 10% to 20% SDS-polyacrylamide gel. The cell lysates were then transferred onto Immobilon-P (Millipore) using the Semidry blotter (Bio-Rad) according to the methods described in the manual. The filter was then blocked at 4°C overnight with 5% skim milk (DIFCO), 2.5% bovine serum albumin (Sigma, A5940), and T-TBS (20 mM Tris-HCl (pH 7.5), 150 mM NaCl, and 0.05% Tween 20). The filter was next reacted at room temperature for 1 hour with the individual rabbit antisera diluted 5,000 fold in the antibody dilution buffer (1% skim milk, 0.5% bovine serum albumin, and T-TBS). It was then reacted at room temperature for 1 hour with the biotin-labeled anti-rabbit Ig diluted 1,000 fold in the antibody dilution buffer. Finally, the filter was further reacted at room temperature for 15 minutes with the horseradish peroxidase (HRP)-labeled streptavidin-biotin complex (Amersham) diluted 2,500 fold in the antibody dilution buffer and washed well with T-TBS. The reacting bands were detected by chemiluminescence using an ECL detection kit (Amersham) according to the manual. Consequently, as shown in Figure 15, a band that reacted with every antibody and is presumed to be derived from the 32-8-1 protein was detected in the liver tissue sample near the 130 kDa.

Example 7 Cloning of an upstream cDNA of 686-1-4 by RACE

The present inventors attempted to obtain a cDNA clone 5′ upstream of 686-1-4, which was cloned from the human heart, by the 5′ Rapid Amplification of cDNA Ends (RACE) method. The details follow.

(1) Cloning of cDNA clone D-2 by RACE

The present inventors performed 5′ Rapid Amplification of cDNA Ends (RACE) by using the Marathon Ready human brain cDNA (Clontech, #7400-1) as the adult human brain cDNA library. The 5′RACE mixture consisted of 5 µl of the Marathon Ready adult human brain cDNA, 1 µl each of 10 µM primer #878 (5′-TTTGTGCCCACCAGAGCCAAGTCAG-3′ (SEQ ID NO: 73)) and 10 µM AP1 primer (which came with the kit: 5′-CCATCCTAATACGACTCACTATAGGGC-3′ (SEQ ID NO: 42)), 1 µl of Advantage™ KlenTaq polymerase mix (Toyobo, CLK8417-1), and 33 µl of deionized water to make the total 50 µl. A PCR reaction using the Thermal Cycler (95°C for 1 minute, five cycles of 95°C for 5 seconds and 72°C for 4 minutes, five cycles of 95°C for 5 seconds and 70°C for 4 minutes, and 25 cycles of 95°C for 5 seconds and 68°C for 4 minutes) did not produce clearly detectable bands. Therefore, the reaction mixture was diluted 50 fold, and 5 µl of this was mixed with 5 µl of 10 x Advantage™ KlenTaq buffer (which came with the kit), 4 µl of 2.5 mM dNTP, 1 µl of 10 µM AP2 primer (which came with the kit; 5′ ACTCACTATAGGGCTCGAGCGGC-3 (SEQ ID NO: 44)), 1 µl of 10 µM 32-8-1 5′ RACE nested primer #757 (5′-GTGAAAGGGGTAAAGGCTTAGCAAC-3′ (SEQ ID NO: 74)), 1 µl of Advantage™ KlenTaq polymerase mix (Toyobo, CLK8417-1), and 33 µl of deionized water to make the total 50 µl. Nested PCR was performed at 95°C for 1 minute, five cycles of 95°C for 5 seconds and 72°C for four minutes, and five cycles of 95°C for 5 seconds and 70°C for 4 minutes. Subsequent treatment of 15 cycles at 95°C for 5 seconds and 68°C for 4 minutes produced a band of 1.8 kb. The products were separated on a 0.8% agarose gel. The corresponding band was excised and purified with the QIAquick gel extraction

kit (QIAGEN, 28706) and subjected to TA cloning according to the manual for the pGEM-T Vector System I (Promega, A3600). The resulting clone was designated D-2.

The nucleotide sequence determination was done as previously described, and the nucleotide sequence of 1,776 base pairs was determined by means of the cycle sequence method using the Dye Terminator Cycle Sequencing FS Ready Reaction kit (Perkin Elmer, Catalog #402122). The nucleotide sequence thus determined was found to encode 590 amino acids (SEQ ID NO: 75).

(2) Cloning of cDNA clone 1.2 kb#33 by RACE

The open reading frame that exists within the sequence of Clone D-2 is a sequence upstream of the 781st nucleotide of SEQ ID NO: 3. This open reading frame is not closed, that is, it does not have a stop codon, so it was assumed that the open reading frame continues further upstream. Thus the present inventors prepared a new primer and performed 5′ RACE. By using the Marathon Ready human brain cDNA (Clontech, #7400-1) as the template, the present inventors performed 5′ Rapid Amplification of cDNA Ends (RACE). The 5′RACE mixture consisted of 5 µl of the Marathon Ready adult human brain cDNA, 1 µl each of 10 µM primer B5R-1 (5′-GCAGATGGAGAACGGGAAACTATGG-3′ (SEQ ID NO: 76)) and 10 mM AP1 primer (which came with the kit; 5′-CCATCCTAATACGACTCACTATAGGGC-3′ (SEQ ID NO: 42)), 5 µl of 10 x Advantage™ KlenTaq buffer (which came with the kit), 4 µl of 2.5 mM dNTP, 1 µl of Advantage™ KlenTaq polymerase mix (Toyobo, CLK8417-1), and 33 µl of deionized water to make a total of 50 µl. A PCR reaction using the Thermal Cycler (95°C for 1 minute, five cycles of 95°C for 5 seconds and 72°C for 3 minutes, and five cycles of 95°C

for 5 seconds and 70°C for 3 minutes, followed by 25 cycles of 95°C for 5 seconds and 68°C for 3 minutes) did not produce detectable bands. Therefore, the reaction mixture was diluted 50 fold, and 5 µl of this was used as the template to perform nested PCR. The template was mixed with 5 µl of 10 x Advantage™ KlenTaq buffer (which came with the kit), 4 µl of 2.5 mM dNTP, 1 µl of 10 µM AP2 primer (which came with the kit; 5'-ACTCACTATAGGGCTCGAGCGGC-3 (SEQ ID NO: 44)), 1 µl of 10 µM primer B5R-2 (5'-GAACGGGAAACTATGGGGCTGACAA-3' (SEQ ID NO: 77)), 1 µl of Advantage™ KlenTaq polymerase mix (Toyobo, CLK8417-1), and 33 µl of deionized water to make the total 50 µl. A reaction consisting of 95°C for 1 minute, five cycles of 95°C for 5 seconds and 72°C for 3 minutes, five cycles of 95°C for 5 seconds and 70°C for 3 minutes, followed by 15 cycles of 95°C for 5 seconds and 68°C for 3 minutes produced a band of 0.8 kb. The products were separated on a 0.8% agarose gel, the corresponding band was excised and purified with the QIAquick gel extraction kit (QIAGEN, 28706). It was then subjected to TA cloning according to the manual for the pGEM-T Vector System I (Promega, A3600). The resulting clone was designated 1.2 kb#33. The results of the nucleotide sequence determination (conducted as described before) revealed that the ATG codon starting with the 71st nucleotide corresponded to the first methionine, and that the clone encoded 235 amino acids. The last amino acid, arginine, corresponded to the arginine encoded by the nucleotides from position 108 to position 110 of clone D-2, and the nucleotides from position 1 to position 110 of the clone D-2 sequence overlapped with clone 1.2 kb#33. Therefore, the present inventors concluded that all the

upstream sequences of the deduced open reading frame were cloned.

(3) Analysis of the difference in amounts of expression in tissues

Amounts of mRNA expression in 24 types of tissues, including those that were examined by the northern blot above, were compared by RT-PCR. Human MTC panel I (K1402-1), human MTC panel II (K1421-1), and human fetal MTC panel I (K1425-1), which are commercially available from Clontech, were used as the cDNA. Results of the PCR reaction with the following reaction conditions are shown in Figure 16-A. PreMixTaq (10 µl; ExTaq TM Version) (Takara, PR003A), 2 µl of 2 µM 686D primer (SEQ ID NO: 66, corresponding to positions 2970 to 2989 of SEQ ID NO: 83, and to positions 1 to 20 of Figure 17), 2 µl of 2 µM 686E primer (SEQ ID NO: 67, corresponding to positions 3635 to 3654 of SEQ ID NO: 83, and to positions 666 to 685 of Figure 17), 1 µl of the first strand cDNA, and 5 µl of deionized water were mixed to make a total of 20 µl and reacted. Reaction of 94°C for 5 minutes, and 30 cycles of a three-step PCR reaction consisting of 94°C for 15 seconds, 50°C for 30 seconds, and 72°C for 30 seconds, followed by an extension reaction of 72°C for 7 minutes were performed.

As a result, high degrees of expression were seen in the brain (1), placenta (7), skeletal muscle (8), ovary (10), spleen (14), testis (15), fetal heart (18), fetal kidney (19), and fetal skeletal muscle (22). In addition, 15 µl of PreMixTaq (ExTaq TM Version) (Takara, PR003A), 3 µl of 2 µM 686D primer (SEQ ID NO: 66), 3 µl of 2 µM XE primer (SEQ ID NO: 15, which corresponds to

positions 3915 to 3934 of SEQ ID NO: 83, and to positions 946 to 965 of Figure 17), 1 µl of the first strand cDNA, and 8 µl of deionized water were mixed to make a total of 30 µl and reacted. Reaction of 94°C for 5 minutes, and 30 cycles of a three-step PCR reaction consisting of 94°C for 15 seconds, 50°C for 30 seconds, and 72°C for 1 minute, followed by an extension reaction of 72°C for 7 minutes were performed. As shown in Figure 16-B, bands of 750 bp, 850 bp, and 950 bp, which are presumed to be derived from three types of transcription products, were detected. Since the experiments in Figure 16-A and 16-B both used the same 686D primer, three types of splicing should take place within the sequences that exist in between the 686E primer and the XE primer, producing the transcription products of different lengths. The three types of PCR products were cloned from those of the fetal heart that showed the highest expression. Cloning was done by excising the corresponding bands and purifying them using the QIAquick gel extraction kit (QIAGEN, 28706), and by following the manual for the pGEM-T Vector System I (Promega, A3600).

(4) Analysis of the gene sequences of clones FH750, FH850, and FH950

The nucleotide sequences for the cloned PCR products were determined according to the method described above. The determined nucleotide sequences of FH750, FH850, and FH950 are shown in SEQ ID NO: 79, 80, and 81, respectively. The sequences of the three kinds of DNA are aligned and shown in Figures 17 and 18. Although the sequences of the three kinds of DNA are identical up to sequence position 731, FH850 diverges from FH950 beginning with the 819th nucleotide, which suggests that splicing takes place at

some sequence immediately preceding this position. As a result of this splicing, FH850 generates a stop codon with the sequence from position 819 to position 821, and the translation is expected to terminate at this position.

In FH750, positions 732 to 941 of the FH950 sequence are spliced out, creating a 210 bp deletion of the gene. However, the protein encoded by the transcription product of the FH750 type splicing is predicted to lack 70 amino acids compared with FH950 type splicing since the sequences before and after the deleted region are expected to be translated in the same frame as in those of FH950.

The sequence obtained by combining 1.2 kb#33 (SEQ ID NO: 78), D-2 (SEQ ID NO: 75), and SEQ ID NO: 3 (Figure 26) was bordered by primer 686D and primer XE derived from the 686-1-4 sequence and was identical to FH750. The clone corresponding to the transcription product expected to be generated by the FH750 type splicing was designated 32-8-1a. (The amino acid sequence of the protein is shown in SEQ ID NO: 82, and the nucleotide sequence of the cDNA is shown in SEQ ID NO: 85.) Clone 32-8-1a can code for 2,000 amino acids. The clone corresponding to the transcription product expected to be generated by the FH950 type splicing was designated 32-8-1b. (The amino acid sequence of the protein is shown in SEQ ID NO: 83, and the nucleotide sequence of the cDNA is shown in SEQ ID NO: 86.) Code 32-8-1b can code for 2,070 amino acids. These two genes possess 13 PDZ domains. Furthermore, the transcription product generated by the FH850 type splicing will contain a stop codon in this region, and it can only code for 1,239 amino acids. This means that it possesses only seven PDZ domains. The clone corresponding to this transcription product was designated 32-8-1c. (The amino acid sequence

of the protein is shown in SEQ ID NO: 84, and the nucleotide sequence of the cDNA is shown in SEQ ID NO: 87.)

Sequence comparisons between "32-8-1b" and "AF00168" (Mus musculus 90RF binding protein 1 (9BP-1) mRNA, partial cds.) are shown in Figure 20; between "32-8-1b" and "AJ001319" (Homo sapiens mRNA for multi PDZ domain protein), in Figure 21; and between "32-8-1b" and "AJ001320" (Rattus norvegius mRNA for multi PDZ domain protein), in Figures 22 through 24. The PDZ domain sequences of the protein encoded by the 32-8-1b gene (SEQ ID NO: 83) are also shown in Figure 25.

(6) Identification of the 32-8-1b high molecular weight protein by western blotting

Human neuroblastoma SH-SY5Y cells and human teratocarcinoma NT-2 cells stimulated by retinoic acid to differentiate into neurons were directly dissolved into SDS-PAGE sample buffer and separated on a 7.5% SDS-polyacrylamide gel by electrophoresis. The proteins were transferred onto Immobilon-P (Millipore) using the Semidry blotter (Bio-Rad) according to the methods described in the manual. The filter was then blocked at 4°C overnight with 5% skim milk (DIFCO), 2.5% bovine serum albumin (Sigma, A5940), and T-TBS (20 mM Tris-HCl (pH 7.5), 150 mM NaCl, and 0.05% Tween 20). It was next reacted at room temperature for 1 hour with the individual rabbit antisera diluted 5,000 fold in the antibody dilution buffer (1% skim milk, 0.5% bovine serum albumin, and T-TBS). It was then reacted at room temperature for 1 hour with the biotin-labeled anti-rabbit Ig diluted 1,000 fold in the antibody dilution buffer. Finally, it was reacted at room temperature for 15 minutes with the horseradish peroxidase

(HRP)-labeled streptavidin-biotin complex (Amersham) diluted 2,500 fold in the antibody dilution buffer and washed well with T-TBS. The reacting bands were then detected by chemiluminescence using an ECL detection kit (Amersham) according to the manual. As shown in Figure 19, a protein whose molecular weight exceeds 250 kDa was detected in both SH-SY5Y and NT-N with either rabbit antisera #1 raised against peptide 32-8-1-17 or rabbit antisera #3D raised against PDZ56 that had been expressed as a GST fusion protein and digested with thrombin to retain only the 32-8-1 gene product. The assumption that the full-length 32-8-1b protein consists of 2,070 amino acids agrees with the molecular weight observed.

Industrial Applicability

By utilizing the proteins and the gene of the present invention, it has become possible to isolate the proteins and their genes that bind to the PDZ domains of the proteins of the present invention. It has been reported that proteins having the PDZ domain interact with the proteins that bind to them and function in the signal transduction related to cell proliferation, cell cycle, malignant conversion, apoptosis, cell adhesion, etc. Therefore, if the relationships between the proteins of the present invention and the proteins that interact with them, as well as the related signal transduction pathways, can be clarified, it should be possible to treat and diagnose disorders related to cell proliferation and others described above by targeting these proteins or their genes. These proteins and their genes are therefore useful for developing therapeutic medicines and diagnostic medicines.

What is claimed is: